Dialog Theory for Critical Argumentation

Douglas Walton

Dialog Theory for Critical Argumentation

Controversies (CVS)

Controversies includes studies in the theory of controversy or any of its salient aspects, studies of the history of controversy forms and their evolution, case-studies of particular historical or current controversies in any field or period, edited collections of documents of a given controversy or a family of related controversies, etc. The series will also act as a forum for 'agenda-setting' debates, where prominent discussants of current controversial issues will take part. Since controversy involves necessarily dialogue, manuscripts focusing exclusively on one position will not be considered.

Editor

Marcelo Dascal
Tel Aviv University

Advisory Board

Volume 5

Dialog Theory for Critical Argumentation
Douglas Walton

Dialog Theory
for Critical Argumentation

Douglas Walton
University of Winnipeg

For Karen with love

Table of contents

Acknowledgements

Acronyms

AAAI	Association for the Advancement of Artificial Intelligence
ACL	agent communication language
AI	artificial intelligence
Arcol	language
ASD	formal dialog system.
ASDDR4	dialog rule
ASDLR5	locution rule
BDI	belief-desire-intention model
CB	formal dialog systems
CBCR1-CBCR6	commitment rules
CBDR1-CBDR3	dialog rules
CBLR1-CBLR4	locution rules
CBV	formal dialog system designed to manage implicit commitments
CBVCK	formal dialog system designed to manage implicit commitments.
CE	formal dialog explanation system
CK	common knowledge
COM	commitment other than common knowledge
CSCA	Computer-Supported Collaborative Argumentation Resource Site
DARPA	Defense Advanced Research Projects Agency
DefLog	defeasible logic system
DFG	Forschergruppe Kommunikatives Verstehen
DMP	defeasible *modus ponens*
FIPA	Foundation for Intelligent Physical gents
HTTP	HyperText Transfer Protocol
IACAS	program for interactive argumentation on a computer
IBM	International Business Machines
INFORM	FIPA communication primitive
KQML	Knowledge Query and Manipulation Language
MAS	multi-agent systems
NYSE-TICKS	New York Stock Exchange-Ticks
OMCS	Open Mind Common Sense System

PPD	permissive persuasion dialog
REQUEST	FIPA communication primitive
RPD	rigorous persuasion dialogs
SE&D	standard of evidence and depth model
SMP	strict *modus ponens*
TRACK	computerized planning system
VAF	value-based argumentation framework
XML	Extensible Markup Language

Introduction

Dialog theory for critical argumentation

Recent developments in argumentation theory and artificial intelligence demand a new look into the foundational notion of dialog around which the new dialectical models of argumentation are being built. Two previous books, *Commitment in Dialogue* (1995) and *The New Dialectic* (1998) postulated several distinctive types of dialog, like persuasion dialog, information-seeking dialog, and deliberation. The view that rational argument is based on a dialog framework is an ancient one that was popular among the Greek philosophers. It was first systematically developed as a field of logic by Aristotle. He called the study of how two parties reason together by asking questions and offering replies and arguments dialectic, using this already well-established term. But the field of study remained largely undeveloped on the periphery of logic (Hamblin, 1970). It pretty much remained dormant as an area for logical research for many centuries until its recent revival.[1] Current systems of dialectical argumentation recognize several distinctive types of dialog seen as conversational settings that provide frameworks for argument use. The motivating idea is to model an argument not just as a designated set of premises and a conclusion, in the style of traditional logic, but as a speech act in which one participant in a dialog puts forward a conclusion as a claim made with a set of premises designed to provide support for that claim.

Dialog theory is beginning to be recognized as important, but there are not many studies on it so far within argumentation. The systems of Hamblin (1970; 1971) and Mackenzie (1971, 1981, 1990) are meant to be used for research on informal fallacies. The dialog systems of Walton and Krabbe (1995) model the argumentative exchanges between two parties in persuasion dialogs, but study other types of dialog as well, and also include consideration of shifts from one type of dialog to another during a sequence of argumentation. The study of dialogs as argumentation frameworks has advanced more quickly in computing in the past ten years, surpassing argumentation theory and beginning to fill the gap. The dialecti-

---

1. Chapter 2 outlines the history of the subject, from its ancient roots to the beginnings of its current revival.

cal conception of an argument as an exchange in a dialog between two parties has proved outstandingly successful for developing argumentation technology in computing, especially in artificial intelligence, by providing an alternative to deductive and inductive logical approaches that can model defeasible reasoning in a more realistic way (Dunne and Bench-Capon, 2006). The emerging field of computational dialectics (Gordon, 1995) takes as its subject matter the study of computational systems of dialog in which two or more parties interact with each other using rational argumentation, as well as explanation, and other kinds of speech acts like the asking of questions. Chapter 1 provides a broad introduction to the place of dialog theory in computing and modern systems of communication.

Especially because of the need to devise systems for electronic communication on the internet, multi-agent computing is moving more and more to a model of argumentation as a dialog between rational agents. Current technologies of artificial intelligence are now widely based on the possibility of communication between entities that can act, reason, ask questions, and exchange information. For example, you might have an agent that filters out your e-mail messages, deletes some, and marks others with a high priority. Or you might have an agent that searches around the internet, and collects certain kinds of information, and then processes it in a format you can use for some purpose. In order to collect this information, the agent will have to ask questions of other agents. Goal-directed communication between agents, or among groups of agents engaged in projects that require teamwork, is more and more important for many applications in electronic commerce and information retrieval. This technology requires a new model of argumentation that fits the foundations of critical analysis and evaluation of arguments found in natural language texts of discourse into a broader setting of dialog in which two or more parties engage in a polite and orderly communication with each other. Chapter 4 explains to the reader how multi-agent dialog systems work in general.

In this book it is argued that argumentation studies need be fitted into a broader field called dialog theory that comprises their critical evaluation. The goal is to provide dialectical methods that can be used to evaluate argumentation as a sequence of conversational exchanges at some stage of a dialog and judge whether the argument actually contributes to the collective conversational goal or not. The goal of dialectic is normative. It is to judge which arguments are stronger and which are weaker (or even fallacious) by appealing to structures based on procedural rules that specify conditions for appropriate uses of an argument. It is meant to utilize not only deductive and inductive forms of argument, but also defeasible argumentation schemes that claim only tentative acceptance for a conclusion, subject to the possibility of defeat as new evidence comes in, or more questions are asked. Such defeasible arguments are useful in conditions of uncertainty where direct knowledge of whether a claim is true or false is not available. Thus they are

very common, for example, in politics and ethics, where situations are highly complex and indeterminate, and a decision has to be made under conditions of changing evidence and lack of knowledge.

The new textbook, *Fundamentals of Critical Argumentation* (2006) shows, using many examples, how dialogs about current subjects of intense controversy can be used to teach students to deconstruct a controversy by analyzing rational arguments found on both sides. The purpose of the present book is not to repeat these interesting and useful examples of controversies formulated as dialogs, but to provide philosophical foundations for the dialog theory needed to contend with them.

As noted above, what is referred to by dialog argumentation, or dialectic, is the idea of two parties reasoning together by taking turns asking questions, offering replies, and putting forward arguments. Such argumentation takes the form of a dialog or conversation between two or more participants. Each participant takes a turn having his or her say. It is a model of thinking in which, as the old expression goes, "two heads are better than one". In this model, an argument is (ideally) a collaborative or polite series of contributions to an orderly conversation (dialog) between two parties. In dialog argumentation, two parties may have a difference of opinions, but are trying to settle it by rational means in a conversational exchange that may be partly adversarial, but is also partly collaborative.

Different types of dialog have been recognized in the argumentation literature so far, and the early part of this book distinguishes these different types and clarifies their relationships. Much has now been written about these different types of dialog, but probably most has been written in the argumentation literature on persuasion dialog. In this type of dialog, one party has doubts about a claim put forward as acceptable by the other party. In such a dialog, there are conflicting points of view on an issue, and each party tries to persuade the other to accept his or her point of view by using rational argumentation. This type of dialog does have an adversarial aspect, and some critics have suggested that it is infected with the pernicious kind of relativism, because it favors acceptance over really finding the truth of the matter being discussed. It seems to follow from this criticism that persuasion dialog should be discounted as more or less a waste of time, because it never really proves anything, and is of no use in finding the truth. One problem posed in this book is how a model of persuasion dialog can be built that meets these objections.

Another problem is how dialogs fit together. Most work in dialog theory has been on the internal structure of a dialog. But there are also important questions about how two dialogs can fit together. In a so-called dialectical shift, argumentation starts out in one type of dialog and moves to another type of dialog. For example, an argument may start as deliberation between two parties and then shift to a negotiation. Two parties may deliberate on how to hang a picture, and then they may begin to negotiate on who should go to get a nail and hammer. An em-

bedding underlies a shift in which the argumentation in the second dialog contributes to that in the first. Failures of embeddings in dialectical shifts in arguments are especially interesting because they are often associated with many of the leading informal fallacies, like fallacies of relevance. The problem investigated in chapter 6 is how to determine whether a shift is based on an embedding or not, in a given case. This problem is approached in a practical and illustrative way by presenting a series of eight examples of argumentation in everyday discourse. In each case there is a dialectical shift; some are embeddings and some are not. Criteria are proposed for solving the problem of dealing with shifts.

Still another problem is how dialectic can be applied to interpreting a text of discourse. If dialog theory is part of a method of critical argumentation that can be used to evaluate arguments as strong or weak, it faces a prior problem. That problem is the notoriously difficult one of interpreting a natural language text of discourse in order to identify the argument to be evaluated, and to analyze the text by justifying one's interpretation and choosing between competing interpretations. The basic problem of the identification, analysis and evaluation of any argument is to try to determine what the premises and the conclusion are supposed to be, as specific propositions that the arguer is making a commitment to having asserted, in the given case. The second problem is, once an argument has been identified, or some parts of it have been identified at any rate, to find any implicit premises in it that may be necessary to take into account before the argument can properly be analyzed and evaluated. The final chapter investigates the fundamental question of how well the dialog theory developed in the book can be applied to the problem of fairly and reasonably interpreting an argument prior to making an attempt to analyze or evaluate it, using the tools now available in critical argumentation.

This book shows that dialog theory has a precise structure that can show how argumentation, and other forms of multiparty reasoning like explanation, can be seen as orderly processes, organized under sets of normative principles and rules appropriate for useful discussions of various kinds. It advances the field by outlining basic principles, surveys the history of it up to its recent revival, and integrates the state of the art with new advances in artificial intelligence and multi-agent computing.

The place of dialog theory

Dialog theory has recently come to be seen as the underlying structure on which to base the analysis and evaluation of argumentation and fallacies. As such, it is now seen as a vitally important part of theory by those of us working in argumentation studies and allied fields. Although it is a very old subject, and leading ancient thinkers made important contributions to it, for two millenia it has lain dormant. Only in the 1970's, notably in the work of Hamblin, Barth and Krabbe, Rescher, Hintikka, and Grice, did the subject reappear in analytical philosophy, although it had been studied earlier by the Erlangen School. But now in the twenty-first century some scientists, mainly computer scientists, are also getting seriously interested in dialog theory, and even seeing it as a necessary part of their research initiatives in fields like expert systems technology and multi-agent systems. There is also growing interest in dialog theory in communication studies (Dascal et al., 2005). Judging from all these developments, dialog theory is finally once again being regarded in the twenty-first century as a subject that is of some importance.

This chapter outlines the philosophical background of dialog theory and briefly surveys some of the latest research initiatives on dialog theory in computer science. The main components of dialog theory are briefly explained. Included is a classification of the main types of dialog that, it is argued, should be the central subject matter of the field. The central problem of dialog theory, retraction of commitments, along with one proposed solution, is briefly explained. Following these surveys, a prediction is made about the direction dialog theory will take in this century in relation to the growing field of communication studies. Some recent developments concern the way dialog theory has assumed a place of growing importance in computer science. Computer scientists are collaborating with argumentation theorists to build new models of reasoning in artificial intelligence that go beyond old models based exclusively on deductive and inductive logic. In these new models argumentation is seen as a dynamic process in which one party puts forward an argument that may change and develop as it is confronted in a dialog with the questions, doubts and criticisms of another party who may or may not accept the argument. This new way of looking at rational argumentation has produced a paradigm shift in the way the cognitive sciences understand rational thinking.

The order of presentation of ideas in this chapter has four parts. The first part briefly sketches out the philosophical and historical development of dialog theory. The second part gives a quick overview of new scientific research initiatives on dialog theory, mainly in computer science. The third part explains, in as simple a fashion as possible, the basic ideas or "building blocks" of dialog theory. The fourth part looks toward the future, showing in brief outline how dialog theory can be expanded in its grasp, to cover not only argumentation theory, but also fundamental notions essential to computing and communication. Each part opens the way for the reader to the further more detailed exposition and development of these ideas in later chapters of the book.

1   The rebirth of dialog theory

Argumentation based on a dialog model has now become widely accepted as an important technology for use in computing, especially in artificial intelligence. At the First International Conference on Computational Models of Argument (COMMA06), held at the University of Liverpool in September, 2006 the extent of this acceptance was clearly stated on the conference web page.[1]

> Over the past decade argumentation has become increasingly important in Artificial Intelligence. It has provided a fruitful way of approaching non-monotonic and defeasible reasoning, deliberation about action, and agent communication scenarios such as negotiation. In application domains such as law, medicine and e-democracy it has come to be seen as an essential part of the reasoning.

In the beginning, argumentation was developed as a tool to help university students think more critically. How from these modest beginnings, did it grow and develop to the point where it has now, in effect, become a subfield of computing? Iyad Rahwan asked me this question at the ECAI (European Conference on Artificial Intelligence) Meeting, in Riva del Garda, Italy, on August 29, 2006, during the Workshop on Computational Models of Natural Argument (the 6th such workshop in the series). This chapter can be seen as an attempt to answer his question. The problem is that the development of argumentation technology has advanced so rapidly that any attempt to publish a survey would be badly outdated by the time it appeared in print. The history of the development of the subject has yet

---

1.   COMMA06 was organized by ASPIC (Argumentation Service Platform with Integrated Components). The Organizing Chairs of the conference were Professors Michael J. Wooldridge, Trevor J. M. Bench-Capon and Paul E. Dunne of the Department of Computer Science of the University of Liverpool. The web page for the conference is: http://www.csc.liv.ac.uk/~comma/

to be chronicled, and this is not the place to attempt that kind of project. But still, as a participant in the process, my answer to the question will be of interest.

My earliest motivation that led to the study of argumentation came from two roots. One was the experience of teaching introductory and advanced logic students at the University of Winnipeg. I quickly came to realize that what they really needed, and seemed to benefit most from, was not additional teaching in propositional and predicate logic, but the development of more practical skills falling into the traditional area of informal fallacies. The only sophisticated treatment of fallacies available at that time was the groundbreaking book *Fallacies* (1970) of Charles Hamblin. Hamblin had put forward formal dialog systems meant to be used for the study of fallacies. The problem was that this was only a first step, and there was still a big gap between the formal dialog systems, which were of a simple kind, and the real (and complex) world of fallacies.

During the years 1975–76, I took part in a logic seminar at Victoria University of Wellington, New Zealand, and during that year I also made a visit to Australia where I was able to talk with Charles Hamblin, but for one day only. In retrospect, it would have been wonderful to spend the whole year studying with him, but that was not the way things worked out. I was to spend my next sabbatical year (1982–83) at the University of Auckland, where I had enough leisure time to work on a monograph applying a new theory of dialogs, essentially based on Hamblin's, to problems in studying the major informal fallacies. Little did I know at the time that this work would take up much of my effort in subsequent years as I addressed each fallacy in turn, often devoting a whole book to a single fallacy. This monograph, called *Logical Dialogue-Games and Fallacies* (1984), appears not to have been widely read, and is still little known.[2] This book laid the foundations for the subsequent research on dialog systems. The book constructed several logical dialog games, basically extensions of Hamblin's systems of formal dialog, and studied fallacies, like the fallacy of circular reasoning (begging the question), in these formal models (Mackenzie, 1979, 1981, 1984).

The book studied strategies in such formal games of dialog. The central idea behind these formal dialogs and the strategies used in them was that the one party, called the proponent, has a particular proposition designated as his thesis to be proved, and she tries to use arguments to get the other party to come to accept this thesis as a commitment. The central idea was that in order for such an argument to be successful, and fulfill the goal of the dialog, it had to not only be structurally correct, but also to be based on premises that are commitments of the respondent

---

2.   For some years now it has been of print, although I have recently put a copy of it on my personal web page for researchers to use (http://io.uwinnipeg.ca/~walton/books/LDG84bk. pdf), as I still think it has some value, although it is long out of date by current standards.

to whom the argument was addressed. This I took to be the central concept of the persuasion dialog, or what I would later come to classify as that type of dialog. I feel that the novelty of the book was its introduction of this notion of persuasion dialog, although I did not appear to use that term in 1984. Still, the essential notion itself was there, because the argumentation in the dialog was centered around this notion that the one party uses the commitments of the other party in order to persuade that second party to come to accept a proposition that he is skeptical about, or doubts. At this stage, like Hamblin, I was still not aware of the distinction that needed to be made between different types of dialog.

Dialog theory was introduced to modern analytical philosophy by Grice, in his ground-breaking paper on the logic of conversation (1975). In the Gricean framework, an argument is viewed as a contribution to a collaborative conversation between two speech partners. From the Gricean point of view, an argument, or any other move (speech act), should be evaluated on the basis of its collaborative value as a contribution to the conversation, at the stage of the conversation where it was put forward. According to Grice, there are so-called conversational maxims that represent guiding rules of polite discourse. These maxims are the basis of implicatures, or inferences suggested to one party by what another party says in the context of a conversation. Although dialog theory was long dormant in philosophy before Grice, it is not altogether new.

Portraying arguments in a dialog framework, in which two parties engage in an exchange of orderly questions and replies, is quite an old idea in philosophy. It was highly familiar to the Greek philosophers, and is best known through the dialogs written by Plato to represent the philosophical activities of Socrates (Robinson, 1953). The art of reasoning in dialog was called 'dialectic' by the Greek philosophers, and Plato and Aristotle, in particular, saw dialectic as the method of philosophy. Aristotle also saw dialectic as fundamental to another closely related field – rhetoric. Dialectic was very important to the Greek philosophers, and Aristotle even made a classification of different kinds of dialog used in dialectical argumentation. As shown in chapter 2, he cited five kinds of arguments used in discussion: didactic, dialectical, examination and contentious (eristic) arguments (*On Sophistical Refutations* 165a38- 165b12). But as will also be shown in chapter 2, the historical development of the notion of dialectic was peculiar. It did survive in some forms into the logic curriculum in the middle ages, but then the dialectical model of argumentation faded into the background.

With the advent of the scientific view of reasoning in the Enlightenment period, dialectical argumentation became an obscure, even alien notion. Especially after the rise of mathematical logic and logical empiricism in the late nineteenth and early twentieth centuries, dialectic was ignored, and had no place at all in logic or the study of rational argument. The Erlangen School in Germany, under

Lorenzen began around 1978 to revive formal dialectic as a subject. The dialog systems of Lorenzen were meant to provide us with viable concepts of notions already familiar from formal logic like logical consequence, logical truth, and logical consistency (Krabbe, 1998, p. 60). For this purpose different systems of formal dialog were constructed that work by assigning a clear dialogical meaning in a context of use to each logical constant (Barth and Krabbe, 1982). Each context of use arises from a conflict of opinions in which a proponent puts forward a thesis and attempts to defend it against challenges put forward by an opponent. Each logical constant is characterized by specific modes of challenge. Several types of dialog systems were constructed by Lorenzen. In one of the most well known of these systems, the opponent begins by challenging the thesis of the proponent. Each subsequent move is then a challenge to answer according to a logical rule. The proponent wins if he successfully defends his thesis against all of the opponent's challenges. Otherwise the opponent wins. There is a limit on the number of dialog moves that is announced at the beginning of the dialog. The Lorenzen dialog logics are outlined in (Barth and Krabbe, 1982) and in (Walton and Krabbe, 1995, pp. 2–4), but full descriptions of technicalities of the systems can be found in (Lorenzen, 1969) and (Lorenzen and Lorenz, 1978) and (Krabbe, 2006).

In the formal systems of dialog devised by Hintikka (1979), each of the two parties has the goal of proving his or her initial thesis, based on the concessions elicited from the other. This basic idea of using the commitments of the other party as the means of proving something is common to both the Hamblin and the Hintikka frameworks. As Hintikka (1979, p. 362) put it, the motivating idea is that the premises allowed to be used by a party in her arguments must either be the other party's initial thesis, or statements elicited as responses from the other party as answers to questions. However, the Hintikka systems has a competitive aspect that is even stronger than in the Hamblin systems. In such Hintikka dialogs (Hintikka, 1979, p. 362), each player keeps a score sheet that is in the form of a *tableau*, or list of statements. The first player to close his tableau, or to fail to give a full answer to a question posed by the other player, "loses" the dialog. The kinds of challenges and responses allowed in Hintikka's rules are often similar to those of Lorenzen. The Hintikka dialog is an adversarial competition in which you try to prove your thesis before the other player proves his. The Hintikka dialog systems are generally less permissive than the Hamblin or Mackenzie ones. In a Hintikka system, a respondent may fail to answer a question, but if he does so, the negation of the presupposition of the question is inserted into his commitment store.

The Rescher system of formal dialog (1977) is also adversarial, although in a somewhat different way from the Hintikka systems. In the Rescher system, each party attempts to lead the other to violate some rule of the game. This adversarial aspect is reminiscent of many of the medieval games, like the obligation game,

where the goal was to trap the other party into making a self-contradictory assertion. Another similarity is that in the Rescher system of dialectic, the roles of the two parties are different. The proponent moves first, and it is her task to make her thesis known and to present arguments for it. The respondent has the more limited role of responding to the prior moves of the opponent. Each party has his/her set of rules. What makes Rescher's system distinctive is his use of the relation of *provisoed assertion*, meaning that one statement is said to usually or ordinarily obtain, provided that another one does, all else being equal. Nowadays, this form of conditional relation would be called a default conditional. Rescher's system was ahead of its time in recognizing defeasible reasoning as central.

Rescher's system of formal dialectic was also novel and original in other ways. Another unusual feature (p. 4) is that the outcome of a disputation is judged by a *determiner*, a third party who not only "presides as a referee", but also "judges over the conduct of the dispute". This feature is in keeping with Rescher's historical motivation for his system, which is to model the procedures of disputation used in the universities in the middle ages (Rescher, 1977, p. 1). According to Rescher, such disputes were presided over by this determiner, or supervising *magister*, who summarized the result of the disputation, and made a ruling on the issue, once the dialog was concluded. Rescher's statement (p. 3) that the study of the process of disputation of the kind he proposes is worthwhile, because it offers "a vivid view of the structure and workings of the validating mechanisms which support our claims to knowledge". It would appear, judging from this remark, that the goal of the Rescher dialog as a structure is to reveal how claims to knowledge are validated. Citing the work of Lorenzen on formal dialog systems, Rescher wrote (p. 73) that his dialog systems are designed to show how a claim to knowledge is validated through revealing the process of reasoning behind it through a sequence of questions and replies between two parties, where one party plays the role of doubter or skeptic.

Charles Hamblin (1970; 1971), the Australian logician, constructed mathematical models of dialog, motivated by the study of the ancient and long-neglected area of fallacies. Independently, Barth and Krabbe (1982) built up their dialectical systems, based on Lorenzen's framework. Other systems already mentioned include those of Hintikka (1979; 1992; 1993), Mackenzie (1981; 1990) and Rescher (1977). The Mackenzie systems (1981; 1990) are patterned on those of Hamblin, but have more precise rules, and are divided into several different systems. These systems, and others mentioned below, were meant to provide formal structures to represent how a sequence of rational argumentation should proceed when one party argues with another in an orderly way. These historical roots of the field of formal dialectic are chronicled in (Walton and Krabbe, 1995), where the field is also developed through the addition of new formal techniques and new formal dialectical systems.

The advent of recent developments in argumentation theory (van Eemeren and Grootendorst, 1984; 1987; 1992), and informal logic (Johnson and Blair, 1985) has given rise to a different way of looking at arguments. For the first time, there has been a systematic attempt to analyze and evaluate everyday arguments, as actually used in real cases in daily conversational exchanges. Instead of the one-liner examples of fallacies being dismissively evaluated without any real attempt to take context into account, serious attempts are being made to grapple with the problems posed by common forms of argument that are fallacious in some cases but reasonable in others. This practical and more realistic approach to the analysis and evaluation of everyday arguments has inevitably broadened the whole notion of rationality (Johnson, 2000). Instead of being seen as a designated set of propositions, an argument is now being seen as a move made in a dialog in which two parties are attempting to reason together.

The monograph *Commitment in Dialogue* (1995) built several formal systems of dialog, but its central aim was to analyze the concept of commitment, the basic concept of formal dialog theory in Hamblin's work on fallacies, and also in my own first book on fallacies (1984). Erik Krabbe and I felt that a dialectical formalization of this notion, and philosophical clarification of the meaning of the concept, would be important contributions to the study of argumentation.

In the introduction to the book, we showed how the notion of commitment in the formal theory of dialog was based on two previous strands in the literature. One of these was of course Hamblin's work. The other was the formal dialog logic of Lorenzen that had been the basis of the book by Barth and Krabbe on dialog logic (1982). Lorenzen's work had been primarily directed towards the formalization of mathematical reasoning, and perhaps for this reason the dialog logic that he developed was not widely taken up by philosophers and logicians. The Erlangen School of dialog logic was not popular outside Germany, and had little influence. The two streams of research were independent of each other[3]. It wasn't until our collaboration on this book that Eric Krabbe and I joined the two streams. As we noted in the book (1995, p. 5), Hamblin's conception of dialectic includes the Lorenzen type of formal dialog logic, but is much broader – Hamblin saw dialectic as a more general study than logic. He thought that logic can be conceived as a set of dialectical conventions of one particular type, but that the concept of a dialectical system is quite general, and could consist of all kinds of conversational exchanges, concerning for example an ordinary conversational exchange of statements about the weather. In

---

3. Hamblin appeared not to be aware of the contribution of the Erlangen school to dialog theory. At least he did not refer to it anywhere in his writings, and when I talked to him he did not indicate any awareness of it or mention it, as I recall. On the other hand, Hamblin had studied mathematics in Germany, and therefore it is unlikely that he would have been completely unaware of the writings of the Erlangen School.

this way, he paved the way for the later development of the pluralistic view that there can be different types of dialog, with different aims and rules.

Although my writing seemed to have little direct impact on the way logic was being taught in philosophy departments, as far as I could tell, it did appear to have quite an impact in the field of speech communication. Researchers in this field were not only interested in using the new developments in argumentation for their own purposes, but it also became clear that some of the tools that they had developed were important in their own right as contributions to the new practical approach to logic. During 1989–90 I was again a Fellow in Residence at NIAS in a group studying fallacies, organized by Frans van Eemeren and Rob Grootendorst of the University of Amsterdam. It became clear to me during these interactions that there was much to be gained by an interdisciplinary approach. However, the history of the subject is there has been antagonism between dialectic and rhetoric since the time of Plato, and that attempting to combine the results of these two fields can only be possible if some of the historical and disciplinary antagonisms between philosophy and rhetoric can be dealt with.

One of the most valuable ideas developed by the Amsterdam School is that of the critical discussion as a framework of argumentation use. This pragmatic and dialectical approach proposed that arguments in a given text of discourse could be analyzed and evaluated in light of the purpose the argument was used for in a conversational setting. Grice (1975) had already set forth this idea by introducing conversational policies to analyze inferences of a kind that he called implicatures. However I found that when I tried to use this notion to teach students of logic and critical thinking how to analyze concepts like relevance, that there was not enough structure there to be helpful for this purpose. But the Amsterdam School presented a set of ten normative rules governing the conduct of rational discussants in the critical discussion model. This was a big move forward, and showed promise in explaining how fallacies can involve violations of procedural rules of a conversation, as well as erroneous or incorrect forms of inference from premises to conclusions of an argument. However, in my opinion it was a limited approach, by itself, because many of the fallacies can only be properly explained and analyzed by postulating different types of dialog, and also dialectical shifts, or movements from one type of dialog to another. In my view the critical discussion represented only one type of dialog, which I classified under the general heading of persuasion dialog, and which can be contrasted with other types of dialog like deliberation and negotiation. Still, the advent of the recognition of the critical discussion as a distinctive type of dialog with clearly stated rules was a big step forward. Even though the critical discussion, as studied by the Amsterdam school, was not a formal model of dialog, it could easily be recognized by students of logic as a distinctive normative framework of argument use.

Another very important contribution from the field of speech communication was the PhD thesis of Arthur Hastings at Northwestern University on argumentation schemes. Argumentation schemes had been known since ancient times. Aristotle wrote a book on them called the *Topics*, and at least two of his other books contained important information about argumentation schemes. Moreover, through the medieval period, leading writers on philosophy, rhetoric and logic tried to employ schemes, both as tools for argument construction, and as methods for argument evaluation. However, the schemes were never developed in a sufficiently precise form to be useful for either purpose. Hastings only gave very simple, and sometimes not very convincing examples, but did explain each of these schemes very clearly, showing how they are meant to be defeasible by using the device of attaching a set of appropriate critical questions to each scheme. The basic idea was that the scheme itself, as representing a form of inference, along with a matching set of critical questions, can be used to create a tool for the identification, analysis and evaluation of arguments found in a given text of discourse in an individual case. This format of using schemes in such a way presented a very powerful tool that could potentially be applied to arguments. There can be deductive and inductive schemes, but the brilliant part of it is that the schemes represent forms of plausible reasoning that are inherently defeasible, like argument from expert opinion. Although useful, the schemes represent forms of argument that are inherently fragile. Not only are they defeasible, but in some instances they are also fallacious, meaning that they can be erroneous in a tricky way, or even used to unfairly get the best of a speech partner in a deceptive way.

To support this new approach to argumentation a dialog theory is in the process of being developed. But how broad or narrow should dialog theory be? How tightly should the notion of a dialog be formalized? What sorts of actual dialogs are meant to be modeled by the theory? Is the theory restricted to dialogs in which two parties are trying to resolve a conflict of opinions by rational argumentation? Or should negotiations, and other types of argumentative exchanges also be covered under the subject of dialog theory? If the scope of dialog theory is wide enough to cover all kinds of communicative exchanges in which two parties are trying to reason with each other, what domains other than critical thinking should it apply to?

The intended use of dialog theory was to provide a normative structure on which to ground methods for improving critical thinking skills, including writing skills and academic research skills, primarily in a university setting. But dialog theory has all sorts of other potential uses. It is now seen as having many important applications in computer science, and this perception has stimulated its growth.

2   Dialog theory in computing

Argumentation and dialog theory have been of interest in expert systems research, especially in so-called "expert critiquing systems" (Silverman, 1992). Critiquing involves "a two-way communication" in which human practitioners and experts can collaborate on problem solving tasks (Silverman, 1992, p. 4). One interesting area of study here is that of cognitive biases, accidents, slips and lapses of various kinds (p. 16). This area of critiquing systems has much in common with the study of informal fallacies. But expert systems research is only one area of computer science that is based on dialog theory. Many other fields of research in computer science are coming to realize the importance of dialog theory as well. Plan recognition has been based on dialog theory in the work of Carberry (1990). And some nice work on using a dialog model to study explanation has been carried out by Moore (1995).

Various AI conferences have advocated the use of dialog theory in computing. One of these conferences even used the expression 'computational dialectics' as early as 1994. This line of research comes under the category of what is now called computational dialectics. Gordon (1994, p. xi) introduced this expression, one that is becoming widely used to describe the application of dialog models in artificial intelligence, as well as other branches of computing.[4] An early conference was the AAAI workshop on computational dialectics at the AAAI-94 Meeting in Seattle in July of that year.[5]

The workshop description defined 'computational dialectics' as meant to describe "an area of activity in AI, which considers the language and protocol of systems that mediate the flow of messages between agents constructing judgment, agreement, or other social choice, to recognize or achieve an outcome in a fair and effective way." The description tells us that dialectic began with the ancients, and is equated by many with rationality. It goes on to say, "dialectic is an idea that simply will not disappear".

Another example of a significant conference on dialog theory was the AAAI Fall 1997 Symposium on Communicative Action in Humans and Machines, held in Cambridge, Mass. This conference re-examined the view of communication as the use of speech acts in dialogs, as contrasted with the view of communication as transmission of information.[6] The symposium considered dialog phenomena ex-

---

4.   Gordon also used the term in two papers (Gordon, 1994 and Gordon, 1996) in which the title even included the term.

5.   The workshop web page may still be found at the following url. http://www.cs.wustl.edu/~loui/comectics.text

6.   See the conference web site: http://www.cs.umd.edu/users/traum/CA/summary .html

tending classical speech act theory, including turn-taking, problem-solving, feed-back, persuasion, and the roles of participants in dialogs. The growing field of expert systems provides a natural application for dialog theory. When a user of advice or information consults an expert source (whether that source is a human or a machine), more is required than the straightforward asking of factual questions. More complex interactive explanations and clarifications are often needed, because the expert may often make statements that the user has all kinds of problems understanding, or being able to implement. Thus sequences of questions and replies – dialog in short – are a vitally important aspect of the implementation of any expert system. These developments not only showed the importance of explanation for designing useful expert systems. They also demonstrated the necessity of seeing the concept of explanation as based on a dialog between two parties.

The notion of using dialogs to model agent architectures in multi-agent reasoning had already been developed as early as 1993, as shown by a paper on cooperative dialogs (Staniford, Bench-Capon and Dunne, 1993). In this paper, dialog graph models were viewed as normative dialog systems containing norms and values used to feature argumentation in which one participant adopts the role of proposer of an assertion, while the other participant adopts an opposition role in which challenges and objections to the proposer's assertion are put forward. Another early paper (Bench-Capon, 1995) stressed the importance of the notion of argument in artificial intelligence and law, building on the Toulmin model. In this paper (p. 5), argumentation is held to offer prospects of real progress in the field of artificial intelligence and law, and is described as a hot topic (p. 12).

Much research in computer science, not only in robotics but in many other areas as well, increasingly uses what is called "agent reasoning" of a kind that is basically the same as what is called practical reasoning, or the Aristotelian practical syllogism, in philosophy. It is goal-directed reasoning by an agent that concludes to a prudent line of action, based on its goals and what it knows about its external situation. An agent is an entity that carries out actions, based on its goals, and that can be aware of information about its external situation as well, including being able to see some of the consequences of its actions. According to Russell and Norvig (1995, p. 652) agents communicate in dialogs by asking each other questions, answering these questions, informing each other about states of the world, making requests or commands to perform actions, making promises, and sharing feelings with each other. Each agent can assume that any other agent it communicates with has goals, and that any such agent will base its actions on these goals. In other words, an agent can assume that another agent will act according to practical reasoning, or what might be called practical rationality. This agent model of practical reasoning is now very well established in computer science. The main concern at present is to extend it to cases of multi-agent reasoning where teams of

agents act together collaboratively. The central problem here is how the agents communicate with each other for the purpose of acing collaboratively.

On May 1, 1999, there was a workshop on agent communication language held in Seattle, preceding the Autonomous Agents '99 Meeting. The subject of the workshop was "specifying and implementing conversation policies". According to the description of the program posted on its website[7], the expansion of agent research is broadening to include the study of goal-directed conversations which fall into several recurrent patterns or types. Conversation policies, to be defined by the conference, are said to be the means of encoding these different conversation types. One of the most important types of dialog that has been studied so far in multi-agent systems is that of negotiation. Negotiation is rightly seen as important, because agents may even have to negotiate on what type of dialog to engage in. There is a project called "Negotiation by Dialectic Argumentation" being carried out by the Queen Mary and Westfield College Electronic Engineering Department of the University of London.[8] This project, built around the notion of an autonomous agent, concerns cases where agents need to come to agreement on a course of action. The predominant mechanism cited as important in the project is that of negotiation. But recent developments have shown that other types of dialog are very important as well for the development of agent technology. Agents need to exchange information. They also need to carry out actions, and hence must be able to reason in a framework of deliberation.

The field of multi-agent systems has grown very rapidly in recent times. The leading problem that needs to be solved is how agents can communicate with each other by exchanging information, by asking questions, and generally by engaging in dialog with other agents. The development of agent communication languages is currently such an important area of research that it should be seen as the leading platform for the development of dialog theory. Also, agent communication technology is being more and more based on argumentation, and is turning to argumentation as the most useful model for agent communication and reasoning.

Another area of research in computing is the study of natural language using computational methods. The Swedish project *s-dime* includes among its objectives the construction of a computational model of dialog moves suitable for modeling a corpus of natural language dialogs (in Swedish).[9] There is also a project at Odense University in Denmark that carries out fundamental research on dialog systems.[10] This research is said to include the theory of cooperative human-machine dialog

7.   http://www.dfki.de/media/workshops/agents99/greaves.txt

8.   http://www.elec.qmw.ac.uk/dai/projects/negot_via_arg.html

9.   http://www.ling.gu.se/research/projects/sdime/sdime_links.html

10.   http://www.mip.ou.dk/nis/research/index.html

and models for interactive speech systems. Both these projects are oriented to pragmatics and linguistics. There is much current interest in linguistics on extending speech act theory to uses of speech acts in wider contexts of communication.

Several Dutch researchers are active in various aspects of dialog modeling in computing. Gerard Vreeswijk has designed an interactive argumentation system called IACAS. It allows a human user to start a dispute, and then engage in interactive argumentation with a computer.[11] It uses a language in which propositions, rules and cases are represented. Vreeswijk (1997) has surveyed a number of formal systems of argumentation. Almost all of these systems include detailed treatment of something that is regarded as highly important in computer modeling of argumentation – the concept of the defeasible inference – a type of inference that is only provisionally acceptable, and that is subject to default as new information comes in to be considered. The counter-examples that defeat a given defeasible inference cannot always be anticipated in advance. Hence defeasible reasoning has been difficult to model using the resources of standard logics, and this problem has led to the possibility of using formal dialog structures to model defeasible argumentation.

Alexy (1989) showed how such dialog systems can be applied to legal argumentation, a program that is now being carried forward by a group of researchers in AI and law including Bench-Capon (1995), Prakken and Sartor (1996; 1998), Verheij (1996; 2000; 2005) and Lodder (1998; 1999).[12] Henry Prakken has been a leader in the research on how logical research on dialog models can be applied to AI and law (Prakken, 1995, 1997, 2001, 2005, 2006). There is quite a body of researchers in the Netherlands interested in dialog theory as applied both to computer science and law. A good idea of this work can be gotten by looking at the journal, *Artificial Intelligence and Law*. (Lodder, 1999) is another good source to get an idea of how dialog structure is being applied to legal argumentation, another topic of recent interest in computer science. Lodder argues for using a dialogical model of legal justification, using structures of dialog adapted from Hamblin, Rescher and Mackenzie. He applies these models to actual legal cases, showing how the argumentation in each case can be modeled as a dialog with participants, commitments, moves and rules of the kind explained in section 3 below. Another important initiative in law and artificial intelligence is the work on defeasible reasoning. Typical legal argumentation is based on rules that are subject to exceptions of one kind or another, giving legal reasoning a certain flexibility that makes it applicable to real cases without falling into a kind of rigidity that would be highly unfortu-

---

11.   You can find it at this web site: http://tcw2.ppsw.rug.nl/~gerard/iacas.html

12.   A survey of recent argumentation methods for artificial intelligence in law is provided in (Walton, 2005).

nate. Thus legal reasoning is often defeasible, meaning that it holds tentatively until new information comes in, but may then fail, or default. Computerized models of defeasible legal argumentation have been advanced by Bart Verheij (1996; 1999). These models support the dialogical view of legal argumentation, and show the importance of dialog theory as applied to legal reasoning.

There is a Computational Dialectics or DFG (Forschergruppe Kommunikatives Verstehen) research group in Germany whose stated goal is to "combine results from AI, in particular nonmonotonic reasoning, and philosophical argumentation theory in order to formalize decision making processes based on dialogs."[13] Their focus is said to be on "models of argumentation contexts which regulate the role of dialog partners, their rights and obligations, burdens of proof etc." Listed as the contact person is Dr. Gerhard Brewka of the Intelligent Systems Department of the University of Leipzig. Within this group a computational dialectics project called Zeno, with Tom Gordon as the leader, has the goal of devising a programming language "for expressing the discourse norms which regulate and coordinate the procedures for making group decisions, acquiring common knowledge, and resolving disputes." The web page for Project Zeno calls computational dialectics a new subfield of Artificial Intelligence.[14]

Another recent development is the Computer-Supported Collaborative Argumentation Resource Site, including the CSCA Discussion List.[15] This site is concerned with research into tools and techniques used to support argumentation in what are called "issue-based information systems" like debate and negotiation.

The Symposium on Argument and Computation, held in Bonskeid House in Perthshire, Scotland, in June-July 2000, brought together a number of researchers in both computer science and argumentation in order to encourage collaboration. At this meeting it became apparent that researchers in the field of artificial intelligence were very interested in using argumentation tools to move that field ahead. Traditionally artificial intelligence had used deductive models of argument, and inductive models as well, but this approach had not proved very successful. One of the biggest problems was that of the defeasibility of the kind of reasoning needed to solve practical problems of artificial intelligence. It was made clear at this meeting that there were two tools of special interest being developed that could be very useful in artificial intelligence. One is the use of defeasible argumentation schemes with attached sets of critical questions. The other is the modeling of argumentation as reasoning used in different contexts of dialog, acknowledging the pluralis-

---

13.   http://pikas.inf.tu-dresden.de/aktivitaeten/ki97-98/FF/cd.html

14.   http://www-fit-ki.gmd.de/projects/zeno.old.html

15.   You can get information about the list here: http://kmi.open.ac.uk/~simonb/csca/index.html

tic notion that the same argument might be used in one way in one type of dialog, but in a different but also in a legitimate way in another type of dialog.

Five interdisciplinary groups worked to produce parts of a handbook on five subjects. One group considered the question of how dialog logics can be used to specify communication protocols. A second group examined the relationship between theories of argumentation and the implementations of defeasible reasoning in computing. A third group examined computational models of legal reasoning. A fourth group studied the relationship between argumentation and computational linguistics. A fifth group investigated computational models of argument.[16] The proceedings of this conference are to be published, providing a document that will explore the foundations of collaborative research between argumentation and computing. Two factors that emerged as very important for multi-agent communication from the conference were argumentation schemes and formal models of dialog.

These early developments will demonstrate to the reader how dialog theory was recognized as useful for computing, early on. Perhaps the best argument for dialectical argumentation is that it is actually being used in computing, and is now widely seen in that field as a leading tool for making computing more efficient by saving costs through development of natural and useful communication technologies. The best place in the field of computing to illustrate the advantages of adopting the dialectical model of argumentation is in multi-agent systems, an area of research outlined in chapter 4. According to Dignum and Greaves (2000, p. 10), "virtually all multi-agents systems employ some type of explicit or implicit conversational layer," but since "theory has lagged practice in this area," work on formal accounts of agent conversation "is in its infancy." Thus chapter 4 is devoted to explaining how the dialog model has been adopted in multi-agent systems.

3   Agent communication

The notion of an autonomous agent is one of a goal-directed entity that can carry out actions by using means-end reasoning. Hence what an agent does, even if it only a software entity, is similar to what humans are doing all the time. This activity could be called deliberation. It is a process of deciding on what to do, based on one's goals and needs, and on the means that appear to be available. But of course, in the course of carrying out actions, an agent will encounter other agents. It is extremely useful for an agent to be able to recognize another agent as also being an agent. This capability is called "plan recognition" in computing (see section 6 be-

---

16. The web site of the conference can be found at this url: http://www.csd.abdn.ac.uk/~tnorman/sac/

low). It is also necessary for agents to deliberate with each other, and sometimes even to negotiate with each other, in order to accomplish their goals. Dignum and Greaves (2000, p. 1) characterize agent communication languages as structures "expressly tailored to support the kinds of collaboration, negotiation, and information transfer required in multi-agent interaction." Thus the necessity for agent communication is fundamental to the development of multi-agent technology. But how do agents communicate with each other, or how can they communicate with each other in a useful fashion so that it is possible for them to collaborate on a task? It seems that agent communication might be a trivial problem. One agent could just ask for information from the other, and then the second agent could pass along that information. No big problem. But actually, even the very simplest example of agent communication reveals that there are a number of subtleties. One agent cannot look directly into the inner mental states of the other. On the other hand, it is necessary to know some things about what the other agent is thinking or intending, in order to have efficient communication that would not be impeded by the asking of all kinds of mundane questions. What an agent must do, in interacting with another agent, is to make assumptions about the thinking of the other agent, and then go ahead and communicate on the basis of these assumptions. If the assumption turns out to be wrong, then the assumption can "default" and be withdrawn.

Two examples of actions involving two agents, called A1 and B0, have been presented by Huhns and Singh (1998, p. 87).

> Suppose agent A1 desires both ice cream and soup, but given that the weather is cold (and based on beliefs not mentioned here), A1 intends to have only soup. Means-end reasoning causes A1 to get soup from the pantry and heat it in the microwave oven. In a second example, A1 sees B0 perform actions or hears B0's statements indicating that B0 is opening the refrigerator door. From this action, A1 uses plan recognition to infer that B0 is about to get ice cream. Knowing B0 to be "rational", A1 figures that B0 does not believe it is cold outside. Since A1 is a helpful agent, he tells B0 that it is cold outside.

In the first example, many of the necessary components of agent means-end reasoning are illustrated. Agent A1 is engaged in deliberation, using means-end reasoning. He scans his environment, realizes it is cold, and decides to have soup. Part of the means to carry out this goal is to heat the soup in the microwave. Another part of the means is to get soup from the pantry. Hence A1 gets the soup from the pantry and heats it in the microwave oven. The second example involves communication between A1 and B0. A1 infers an assumption based on his observations of B0's external actions. He sees or hears something indicating that B0 is opening the refrigerator door. From this data, A1 draws a number of conclusions by infer-

ence. He concludes that B0 is (presumably) about to get ice cream to eat. From this conclusion he infers yet another conclusion, also a guess, but a reasonable assumption. He infers that B0 does not believe it is cold outside. How does he arrive at these conclusions? This is not an easy question. It seems that he arrives at them because he can assume that B0's thinking is pretty similar to his own in many ways. He also assumes that many normal ways of doing things are known to both himself and B0, and are the basis of their actions. For example, he knows that the ice cream is normally kept in the fridge, while the soup is normally kept in the pantry. He also knows, or assumes, that it is more pleasant to eat soup rather than ice cream on a cold day. By using all these normal assumptions, A1 can infer yet another conclusion. He can draw the conclusion that B0 does not know that it is cold outside. Now it would be pointless to tell B0 that it is cold outside if B0 already knows this. But it is very helpful to tell B0 this information if B0 does not know it. This case is similar in some ways to the kinds of cases studied by Grice (1975). In one of these famous cases (see chapter 4, section two), a professor reads a letter of reference written by a second professor on behalf of a student who is applying for a job (Grice, 1975, p. 163). The letter praises the student's command of English and his class attendance record, but says nothing else. What is the significance of what has been left out? The second professor would certainly draw a conclusion from this omission, by a process of assuming that the letter is meant to convey a negative message. The case is similar to that of A1 and B0. In both cases, one communicator is assuming that the other probably knows or does not know certain things, and both are aware that there are familiar ways of doing things, like eating food or writing a letter of reference. The inferences drawn are based on these background assumptions.

These apparently simple cases show how communication between agents is a lot more complicated than you might initially think, before actually trying to build an agent that can communicate with other agents in a practical way. There is a lot of reasoning going on, in the form of drawing conclusions. But the reasoning, or argumentation, is not made up of deductive inferences. Instead, it is guessing or drawing of assumptions as tentative conclusions that are useful to act on, but that might turn out to be wrong. The one agent is making a lot of assumptions about what it takes to be the internal states and private thinking of the other agent. The one agent is using what might be called empathy, or is sometimes called simulative reasoning (Barnden, 1995). It assumes that the other agent thinks roughly the same way it does, that certain feature of the environment are "normal", and that both agents take this normality for granted. The possibility of communication is based on background assumptions to the effect that in a collaborative conversation, both parties in a dialog will operate on certain unstated assumptions about what the other knows (presumably) or does not know. This feature makes dialog

theory more complicated than one might have initially anticipated, because the principles making for efficient communication are not explicitly stated anywhere. Each party simply assumes that the other party is acting and thinking, at least in general outline, in roughly the same way as she herself would think or act in certain kinds of situations.

The basic components of agent communication can be displayed in figure 1.1 below.
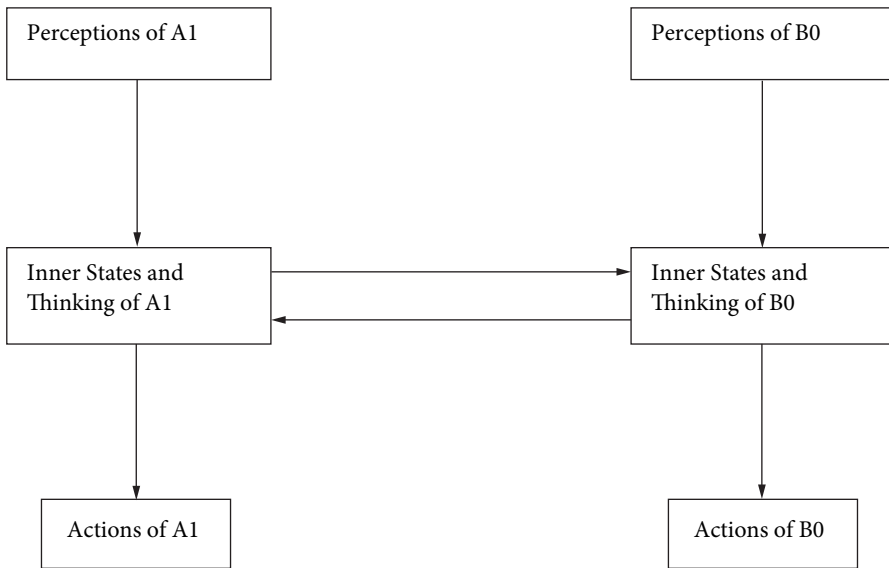
Figure 1.1 Agent Communication

Figure 1.1 shows how one agent is basing its deliberations on how another agent is presumably thinking. Of course, the one agent cannot see directly into the mental states of the other. But it can draw conclusions about these things based on how it would expect the other agent to act in a normal situation that both are generally familiar with. These cases show that the same kind of problem of understanding dialog is present whether the communication is spoken or written. In a written message, like the Gricean case, there is not a dialog in the sense that each party takes turns verbally replying to the previous remark of the other party. But it will be argued in this book (especially in chapter nine) that even argumentation in written discourse can be analyzed and evaluated on the basis of dialog theory.

Even the very simple case of verbal agent communication presented by Huhns and Singh, which shows what appears to be a very mundane and ordinary case of

agent communication, has elements that could be hard for dialog theory to model. It is true that agent A1 and agent B0 are engaging in dialog of a sort. But much of the basis of how they communicate is unstated, and is based on assumptions and expectations that we normally take for granted, but that could be hard to specify using traditional logic. It looks like the agents are doing a lot of guessing based on empathy. And yet there is a dialog. The one agent is presuming that the other does not know something that the first agent does know. He is not sure about this, but it is a good guess, and the usefulness of the information he passes on to the other agent is based on this assumption. The communication is grounded on what is often called a lack-of-evidence inference. Agent A1 does not know whether agent B0 knows that it is cold outside. But he draws the inference by presumption, after observing B0's actions, that B0 does not know this. In other words, there is no evidence that B0 does know this. Therefore, drawing a conclusion from this lack of evidence, A1 infers that B0 does not know that it is cold outside. This form of argumentation is now often called abduction, or inference to the best explanation of a given set of data (Josephson and Josephson, 1994).

The systems of dialog theory that have been constructed so far as formal dialectical models are quite simple. There are two participants, and they take turns making moves. These moves take the form of a dialog. Certain types of moves, and sequences of moves, can then be studied in certain types of dialog. This much of dialog theory is fairly unproblematic, given the current tools and methods. And it can tell us a lot about argumentation, helping us to formally model patterns of argumentation associated with the traditional informal fallacies. But the idea of one participant basing its moves on conjectures about what the other participant is (presumably) thinking is more opaque. The problem is that even in everyday life, we cannot really know about the intentions and beliefs of another person, and are often wrong about such matters. The problem then is how dialog theory can get past this difficulty. It will be argued in chapter 6 that there is a way to do it. But to get to that point, it is necessary to learn more about the existing resources and methods of dialog theory.

For the present, it is easy to see from this small case that multi-agent technology is based on the possibility of agents being engaged in intelligent deliberations with each other. It is also not too hard to see how such deliberation, if it is to be intelligent, must be based on incoming information that an agent can receive. The agent must not only be able to understand this information, and act on it. He must be able to communicate it to others. Thus information-seeking dialog also seems to be an important part of agent communication. Yet another type of dialog that has been shown to be important in multi-agent systems is negotiation. For example Krothapalli and Deshmukh (1999) have shown how inter-agent negotiation mechanisms are necessary for multi-agent manufacturing systems. Hierarchical

manufacturing systems, in which a central controller plans all sequences of opera-
tions, lack flexibility and are too complex. A better alternative is a decentralized
architecture composed of autonomous agents. To make such an operation work,
however, agents have to communicate and negotiate with each other (Krothapalli
and Deshmukh, 1999, p. 1602). Clearly then, in order to perform group tasks, ne-
gotiation between agents is a necessary form of dialog. And yet it would appear
that negotiation is a different type of dialog from either deliberation or informa-
tion-seeking dialog.

The fact that at least three distinct types of dialog are necessary for the devel-
opment of multi-agent technology raises some basic questions for dialog theory. It
suggests that there should be various different types of dialog with different goals
and different kinds of communication methods. What is suggested is a kind of
pluralism. But underlying this pluralism, it would appear that each type of dialog,
as dialog, will also share certain features and characteristics.

4   Fundamental concepts of dialog theory

The simplest illustration of a dialog is a case where one party asks another party a
question. Even if the second party fails to offer any response, the context may indicate
that some relevant reply is called for. Hence here we have a dialog. Not only that, the
very asking of the question, in the right context, has normative implications. Certain
responses do not count as an answer. Certain responses do not even qualify as a rel-
evant reply. Thus even an apparently unstructured case of dialog in which someone
asks a question in a casual conversation may impose enough normative requirements
that it is useful to see it as a dialog in the sense important for dialog theory. So a dialog
is a verbal exchange between two parties, according to some kind of rules, conven-
tions or expectations. But how precise and explicit do the rules need to be?

To answer this question without confusing ambiguity, a distinction needs to
be made between the descriptive study of dialog and the formal study of dialog
(Hamblin, 1970, p. 256). The descriptive study of dialog is concerned with actual
conversational exchanges like parliamentary debates, legal cross-examinations,
and so forth. The formal study of dialog "consists in the setting up of simple sys-
tems of precise but not necessarily realistic rules, and the plotting out of the prop-
erties of the dialogs that might be played out in accordance with them" (p. 256). A
dialog in the formal sense is an exchange with a certain kind of structure – that of
the so-called formal model of dialog. In actual dialogs it is not always clear what
the rules are. In a formal dialog, certain kinds of rules are laid down precisely. The
value of the formal dialog is that it can be applied to an actual dialog in a given
case, and used as a tool to help analyze the case.

Following the general outline of Hamblin (1970; 1971) and Walton and Krabbe (1995), four fundamental building blocks of any formal dialectical system can be identified: (1) the two participants, called the proponent and the respondent; (2) the types of moves (taking the form of various speech acts) that the two participants are allowed to make, as each takes his or her turn to speak; (3) the sequence of moves, in which the appropriateness of each move depends on the type of preceding move made by the other party; (4) the goal of the dialog as a whole. The sequence of moves should ideally move towards the fulfillment of the goal as the dialog proceeds. In any case of actual dialog, one can evaluate arguments and other moves made in the dialog, according to whether or not they are productive, at the stage they were made, in moving the dialog towards its ultimate goal.

This outline gives the general reader a basic structure of the nuts and bolts of dialog theory, following the way the basic concepts and principles were set out by Hamblin. The literature in the field has developed many different formal systems of dialog for various purposes. But they are all based on the basic building blocks clearly set out in Hamblin's account of the structure of formal dialectic as a systematic field. Later, in chapter 6, various limitations in the basic structures outlined by Hamblin will be discussed. In that chapter, Hamblin's basic system of dialog theory will again be reviewed, to bring out certain special features that will be the basis of the discussion in chapter 6. Thus the reader could look ahead to chapter 6 to get more details of the basic components of formal dialectic, and to see how they are put together.

Various formal tools used in dialog theory have proved to be useful for this purpose. One that has been long established is the tableau method. The tableau method of modeling the ordered sequence of moves in a dialog in two columns has been used by Rescher (1977), Hintikka (1979; 1992; 1993; 1995), Hintikka and Hintikka (1982), Barth and Krabbe (1982) and by Carlson (1983). A simple illustration of how the method works is given in Table 1.1. The letters A, B, C,…, stand for propositions. In this illustration, the notation used is similar to that of Hamblin (1971). Why-questions are allowed, as well as various other sorts of moves that indicate whether a participant accepts a particular proposition or not.

Table 1.1 Sample profile tableau of dialog

Proponent	Respondent
1. Why should I accept A?	Because B, and if B then A.
2. Why should I accept B?	Because you accepted it before.
3. All right, I accept B.	Do you accept 'If B then A'?
4. Yes.	Do you accept A?
5. No.	You are inconsistent!

In the sample dialog illustrated in table 1.1, the proponent starts the sequence by asking the respondent why he she should accept proposition *A*. The respondent replies with a deductively valid argument that has *A* as its conclusion. What is illustrated here are two kinds of moves explained below. One is the why-question, which asks for a justification in the form of an argument. The other is the putting forward of an argument, a kind of move in which one proposition (or a set of them) is cited as a basis for support of a particular proposition. The dialog proceeds as the proponent continues the sequence of argumentation with the respondent. At the last move in the sequence, the respondent shows how the proponent has committed herself to an inconsistency. The illustration in table 1.1 shows how the tableau method can be used to represent the sequence of argumentation in a dialog exchange. Each number at the left represents a *round*, or pair of moves by one party and then the other The dialog begins with the opening move, and then each pair of move-numbers represents a so-called "round" or adjacency pair (Hamblin, 1970, 1971; Mackenzie, 1981, 1990).

Four kinds of moves are especially important in dialectical systems: (1) the asking of questions, (2) the making of assertions, (3) the retracting of assertions, and (4) the putting forward of arguments. An assertion contains a proposition, and a form of speech indicating the assertor is committing herself to that proposition in a strong way, implying she is willing to defend it if challenged to do so. Two fundamental types of questions are yes-no questions and why-questions (Kestler, 1982; Walton, 1989). A *yes-no question* admits of only two direct answers – the affirmative answer (yes), and the negative answer (no). A yes-no question is designed to rule out the option "I don't know" as an answer or acceptable reply. The yes-no question is typical of what is called a *choice question*, which limits the options in any allowable direct answer to a definite set of choices. The *search question* does not restrict the respondent to some definite set of choices as direct answers. A narrative answer that tells a story could even be a direct answer appropriate for a search question. The questioner who asks a choice question is more controlling than the questioner who asks a search question, leaving the respondent less freedom to offer any information he thinks may be relevant.

The formal dialectical structures presented above are meant to model argumentation. But it is quite possible that such structures could be expanded to model other kinds of speech acts as well. For example, formal dialectical structures could be used to model different kinds of explanations. Such an expansion of the scope of formal dialectic would be a highly attractive prospect from the point of view of communication theory. Primarily however, so far dialog theory has been concerned with arguments, and with the task of evaluating arguments as strong or weak, correct or incorrect, reasonable or fallacious. For this purpose, its primary use has been to evaluate a given argument from a normative point of view, which

could also be called the point of view of critical thinking. The aim is to spot the weak points where an argument could be criticized as falling short of the standards needed to make it strong, so that acceptance of the premises provides good reasons for the rational acceptance of the conclusion.

The Gricean framework of conversation posed a problem that Grice did not solve. The problem is that there can be different types of conversational exchanges. But how can we classify these different types of exchanges in terms of their goals and rules? A solution to this problem has now been presented in *The New Dialectic* (1998). The new dialectic classifies many different types of dialog that represent different kinds of goal-directed conversations in which argumentation is used to contribute to the goal of the dialog. Six basic types of dialog are described in the new dialectic – persuasion dialog, the inquiry, negotiation dialog, information-seeking dialog, deliberation, and eristic dialog. The properties of these six types of dialog are summarized in Table 1.2.

Table 1.2 Six basic types of dialog

TYPE OF DIALOG	INITIAL SITUATION	PARTICIPANT'S GOAL	GOAL OF DIALOG
Persuasion	Conflict of Opinions	Persuade Other Party	Resolve or Clarify Issue
Inquiry	Need to Have Proof	Find and Verify Evidence	Prove (Disprove) Hypothesis
Negotiation	Conflict of Interests	Get What You Most Want	Reasonable Settlement Both Can Live With
Information-Seeking	Need Information	Acquire or Give Information	Exchange Information
Deliberation	Dilemma or Practical Choice	Co-ordinate Goals and Actions	Decide Best Available Course of Action
Eristic	Personal Conflict	Verbally Hit Out at Opponent	Reveal Deeper Basis of Conflict

In the new dialectic, each type of dialog is used as a normative model that provides the standards for judging how a given argument should be correctly used in a given case. The assumption is that the text of discourse in the given case will provide enough evidence to warrant viewing the argument as supposedly being meant to be a contribution to that type of dialog. In a deliberation dialog, the goal is for the participants to arrive at a decision on what to do, given the need to take action. Hitchcock, McBurney and Parsons (2001) set out a formal model of deliberation dialog in which participants make proposals and counter-proposals on what to do.

In this model (p. 5), the need to take action is expressed in the form of a governing question like, "How should we respond to the prospect of global warming?" Deliberation dialog may be contrasted with persuasion dialog. In persuasion dialog, the one party, called the proponent, has a particular thesis to be proved, while the other party, called the respondent, has the role of casting doubt on that thesis. The goal of the proponent is to prove the proposition that is designated in advance as her ultimate thesis. In the first type of dialog, called the persuasion dialog, the one party, called the proponent, has a particular proposition designated as her thesis, and her goal is to prove this proposition by means of the kinds of argument accepted as persuasive in the dialog. One special type of persuasion dialog is called the critical discussion. This type of dialog will be explained more fully in the next section. The basic feature of argumentation in all the types of dialog is that the one party takes the commitments of the other as premises in arguments. Then by a series of steps of inference, this party uses these premises in argumentation that aims towards providing reasons to support the ultimate conclusion, making it acceptable to the other party.

The more standard account of reasoning in philosophy and the social sciences today is based on knowledge and belief. Belief is taken as referring to the internal mental state of a thinker. Knowledge is taken to be a special kind of belief that is true and justified. But the dialog structure outlined above is based on a participant's commitments, as opposed to his actual beliefs. Another word for commitment is acceptance. 'Commitment' refers to what an arguer has gone on record as accepting, as far as one can tell, from what she has said (and/or done) in a given case, according to the evidence provided by the text and context of discourse in the case. The reason for preferring commitment over belief is that the current work in argumentation takes as its primary goal the evaluation of a given argument, based on the given text of discourse in which the argument was put forward. For this purpose, the arguer's actual beliefs may be extremely difficult to determine, and judging them is a psychological task. On the other hand, his commitments can be more easily determined judging from what he has said (using the given text of discourse). So the model of dialog outlined here, and preferred in argumentation theory, is commitment-based. In the future however, this model can be extended to model belief, knowledge, and intentions.

According to the theory of Walton and Krabbe (1995) there can be dialectical shifts, or changes of context from one type of dialog to another during the same continuous sequence of argumentation. For example, a contractor and a homeowner may be engaged in negotiation dialog on a contract to install a concrete basement in a house. But then the argumentation may shift to a deliberation on whether installing a new concrete basement is a good idea, or whether some other alternative would be better. Or at another point, the discussion may shift to an

information-seeking dialog, in which the contractor informs the homeowner about the city regulations on thickness of concrete for house basements. Functional embeddings of dialogs of this sort have been discussed as a problem for computer modeling by Reed (1998). Reed (p. 250) uses an example of two agents deliberating on how to do a job, and then they begin negotiations when one proposes hiring the other to do the job. In practice, there can be many different kinds of dialectical shifts of this kind. In some cases, the new dialog contributes to the success of the previous one. In other cases, the one dialog is simply an interruption in the progress of the first one. But then the first dialog can be easily resumed. But in some cases, the advent of the second dialog blocks the progress of the first one, and presents a serious obstacle to its progress. The problem of how to formally represent such functional embeddings of dialogs has not yet been solved. It is by no means a purely philosophical problem, and also represents a real problem for the development of computer dialog systems – for example, in multi-agent dialog systems. Chapter 8 is devoted to a presentation of this problem, using several examples (cases) of dialectical shifts in everyday argumentation to illustrate the phenomenon in a practical way.

5   The critical discussion as a type of dialog

The critical discussion is not itself a formal model of dialog, but is the distinct type of dialog that has been most recognized and studied within argumentation theory. Most readers unfamiliar with dialog theory or argumentation theory can easily recognize the critical discussion as something they are familiar with. And indeed, for many people, the critical discussion represents the only type of dialog that they would say is a context of conversation in which genuine argumentation properly takes place. Certainly the critical discussion has a central place among the types of dialog important for the study of argumentation. And thus it is useful to summarize the main features of the critical discussion as a type of dialog here.

The goal of a critical discussion is to resolve a conflict of opinions by means of rational argumentation. There are two participants in the simplest case, called the proponent and the respondent, and each has a designated proposition or thesis to be proved. The two are in competition with each other, for the thesis of the one is the opposite or negation of the thesis of the other. That is what is meant by conflict of opinions, at least in the case of a dispute, a symmetrical type of dialog. However there is another type of conflict of opinions in which only the proponent has a positive thesis to be proved, and all the respondent needs to do in order to be successful in resolving the conflict in his favor is to cast doubt on the proponent's attempts to prove her thesis. This is an asymmetrical type of dialog called a dissent.

Thus it is often said that what distinguishes the two types of critical discussion is the difference in the burden of proof. In the dispute, each has a positive burden of proof. In the dissent, the proponent has a positive burden of proof while the respondent has no burden of proof.

According to the analysis of argumentation of van Eemeren and Grootendorst (1992, p. 35) a critical discussion has four characteristic stages. In the confrontation stage, the two parties agree on the issue to be resolved by the discussion. Each party has a viewpoint, meaning a proposition and an attitude of pro or contra toward that proposition. The viewpoint of the one is opposed to that of the other, forming a conflict of opinions, or "issue" as it is sometimes called. The goal of the critical discussion is to resolve this conflict of opinions by means of rational argumentation. The second stage is the opening stage, where each party accepts procedural and material starting points and agrees to abide by them. The third stage is the argumentation stage, where each side puts forward arguments to support its viewpoint, and attack the arguments put forward by the opposing side. The purpose of this stage is to test the justifiability of the standpoints previously expressed by both parties to the discussion at the confrontation stage. Fourth, there is a concluding (closing) stage that determines the outcome of the critical discussion, by deciding which side is the winner and which is the loser. The participant who has the weaker argument must concede to the party who has the stronger one. This outcome resolves the conflict of opinions. This analysis into the four stages is true of the critical discussion, but it can be applied to all types of dialog in which argumentation takes place. The important things are that argumentation is seen as an orderly process that goes through four ordered stages and moves through these stages aiming toward a collective goal. Thus while the argumentation stage is highly adversarial, the process of rational argumentation can be judged as an objectively determined outcome, based on the normative model, and the rules governing it.

In real argumentation the actual course of a discussion need not temporally follow the order of the four stages. The four stages are ideal dialectical structures representing how argumentation should be viewed normatively. At the confrontation stage the conflict of opinions to be resolved is clarified and identified. To do this, the so-called viewpoint (point of view) of the one party needs to be identified, and there has to be some expression of doubt or disagreement by the other party. At the opening stage, the parties come to an agreement to resolve this conflict of opinions by engaging in rational argumentation. At the argumentation stage, one party takes on the role of proponent (protagonist) and the other party takes on the role of respondent (antagonist, opponent). Each side then puts forward arguments to support its viewpoint. These arguments are supposed to be of certain types, having a particular form or structure. So-called argumentation schemes represent the proper forms arguments should take. Complex argumentation sequences are built

up of simple steps, each of which fits an argumentation scheme. At the concluding stage, the two parties come to a common evaluation of the outcome of the dialog. If the proponent has successfully defended her point of view, then the conflict is resolved in her favor. Otherwise, the conflict is resolved in favor of the respondent. For the goal of the critical discussion to be fulfilled, one or the other of these outcomes must occur.

The argumentation stage of the critical discussion is governed by ten rules (van Eemeren and Grootendorst, 1987, pp. 184–293). These ten rules can be expressed in a somewhat simplified form as follows: (1) Parties must not prevent each other from advancing arguments. (2) An arguer must defend her argument if asked to do so. (3) An attack on an arguer's position must relate to that position (and not some other position). (4) A claim can only be defended by giving relevant arguments for it. (5) An arguer can be held to his implicit premises. Rules (6) and (7) can be combined into a single requirement expressed by the following condition: an argument must be regarded as conclusively defended if its conclusion has been inferred by a structurally correct form of inference from premises that have been accepted by both parties at the outset of the discussion. (8) Arguments must be valid, or be capable of being made valid by the addition of implicit premises. (10) Formulations must not be unduly vague or ambiguous. It can easily be appreciated how violations of these rules correspond to, or at least are closely related to various informal fallacies that have long been recognized as important in logic. For example, committing the *ad baculum* fallacy could easily be identified as a type of violation of Rule (1), the rule that forbids one party from trying to use force to try to prevent the other party from advancing arguments. Fallacies relating to burden of proof, like the fallacy of appeal to ignorance or the fallacy of begging the question, could be violations of rule (2). Rule (3) obviously concerns the straw man fallacy of misrepresenting an opponent's commitments to make his argument look weaker, and more easily refutable. Rule (4) concerns the whole area of relevance. Fallacies that concern failures of relevance include *ignoratio elenchi* (missing the point or "wrong conclusion"), *ad hominem*, *ad populum*, and *ad baculum* (although possible many other known fallacies could come under this heading as well). The fact that *ad baculum* could be either a violation of rule (1) or rule (4) suggests that (perhaps unfortunately) there is no exact one-to-one correspondence between occurrences of specific fallacies and occurrences of critical discussion rule-violations.

The critical discussion is not a formal model of dialog. And it is perhaps because of its realistic and practical nature as a model that it is so easy to recognize it immediately as a common framework of argumentation usage. Once the reader grasps the four stages and ten rules of the critical discussion, it is easy to recognize that the critical discussion can function as a normative model representing how a

productive discussion should ideally use argumentation to resolve a conflict of opinions. Thus one can easily appreciate how the model has normative bite. It is a requirement of the model that the parties agree not only to engage in dialog, but also agree to take part in the critical discussion and to properly take on the roles to be played by each participant. Since the parties are bound by these agreements through all stages of the critical discussion, it would be inappropriate for them to start attacking each other in a quarrelsome way, characteristic of eristic dialog, no matter how tempting this might be. Thus a crowd-pleasing argument that is actually persuasive to a particular audience would not necessarily be a useful or winning argument in a critical discussion. To have a productive critical discussion, it is clear that it is necessary to avoid the various counter-productive tactics associated with the various informal fallacies. Even so, the critical discussion could be an extremely useful model for studying actual cases containing argumentation falling under the headings of the various fallacies. The critical discussion could be used as a normative model to analyze how such persuasive argumentation tactics are successfully used to deceive the inattentive or uncritical arguer. On the other hand, it can also be useful to investigate how the critical discussion could be formalized. If it could be precisely expressed as a formal model of dialog, the outcome could be extremely useful, both in logic and in computing as well.

Perhaps unfortunately however, it seems that there is no single formal model that can capture all of the characteristics of the critical discussion in one precise but reasonably simple structure. There seem to be several reasons for this failure. One is that the simplest dialog structures that seem to represent logical argumentation are quite rigid. They do not allow for retraction of commitment, and they do not seem to represent realistic argumentation in many ways, which is more flexible and open. Another observation is that many real dialogs seem to be quite successful even though the original conflict of opinions was not resolved by showing one party to have successfully won the discussion by persuading the other party to give up his former viewpoint. Take a philosophical discussion, for example, of the kind typical of the Platonic dialogs. Particularly in the early dialogs, like the *Meno*, two opposing viewpoints are put forward, and many arguments are expressed on both sides. But there is no clear winner or loser. The dialog ends in a draw. Even so, the argumentation in the dialog can be quite enlightening. It can have a so-called maieutic effect of bringing new ideas to birth, by critically examining a received opinion and subjecting it to probing scrutiny. It seems in such a case that there has been a persuasion type of dialog, and a successful one at that, at least from a maieutic point of view, even though the original conflict of opinions was not resolved in the end. For all these reasons, a distinction is made in (Walton and Krabbe, 1995) between persuasion dialog and critical discussion. This distinction will be further clarified in chapter 3, where the maieutic effect will be analyzed more fully.

Very often the terms 'persuasion dialog' and 'critical discussion' are used inter-changeably, and in fact this convention can often be followed without any harm or confusion. Nevertheless for purposes of dialog theory, it is best to clarify this rela-tionship by treating the critical discussion as a special type of persuasion dialog. Persuasion dialog is the more general category in which the aim of the one party is to persuade the other party. 'Persuasion' in this sense refers not to psychological persuasion but to a kind of rational persuasion. A proponent in a dialog persuades a respondent to accept a particular statement in this special sense when the propo-nent presents an argument containing only premises that the respondent is com-mitted to, and uses this argument to get the respondent to become committed to the conclusion of the argument. Presumably, the respondent was not committed to this statement previously. So persuasion, in this sense, refers to the respondent's "conversion" so to speak, or the change in his commitments. Before he was not committed to this particular statement, but now he is. In (Walton and Krabbe, 1995) many different formal models of persuasion dialog are constructed and stud-ied. Some are comparatively simple, and are purely formal types of dialog. These are called rigorous persuasion dialogs or RPD's. These RPD's have the advantage of being precise and formally rigorous. They have a precise mathematical structure, and a computer could easily be programmed to engage in this type of dialog, either with another computer or with a human user. But the RPD's do not model realistic natural language argumentation very fully, because natural language persuasion dialog is more flexible and open in certain ways. In an RPD only certain simple moves are allowed at each turn, and only very simple types of responses to each move are allowed. In a PPD (permissive persuasion dialog), a participant can make several kinds of speech actions in a given move, and the ways the respondent can reply are also more variable. This flexibility is more characteristic of natural lan-guage persuasive discourse, but it is also harder to model in a relatively simple and precise dialog structure. The key problem in all such attempts to construct a formal model of dialog is that of retraction.

6   Plan recognition and deliberation

As noted above, research in AI based around the idea of agents performing practi-cal tasks depends on the underlying assumption that agents need to communicate in order to carry out ordinary actions. This need to structure agent communica-tion has proved to be especially vital in the field of AI called planning. As indi-cated by the simple examples discussed above, for agents to collaborate on a plan, a capability called plan recognition is also necessary. One agent needs to be able to interpret the actions of another agent as being based on some kind of coherent

means-end reasoning directed towards a goal. Much research on plan recognition technology has now been carried out, and many software systems have been developed. A nice survey of the historical development of plan recognition technology has been presented by Carberry (1990). A brief summary of some of the highlights of this survey can show how the technology works and how the concept of dialog is at its basis. According to Carberry (1990, p. 17), plan recognition is the method whereby "an agent attempts to reconstruct from the available evidence a plan that was previously constructed by another agent." A simple example from (Carberry, 1990, p. 17) illustrates the kind of problem posed for plan recognition. A driver sees an empty car parked on the highway. She also sees that the car has a missing tire. Further down the highway, she sees a man rolling a tire along the roadside. He is carrying a baby as he rolls the tire, and three small children are following behind. Just by observing these facts, the driver draws several inferences. She infers that this man rolling the tire was the driver of the car she saw further down the highway with the missing tire. She infers that the man was rolling the tire along in order to get it repaired. She infers that the man did not want to leave the three small children alone in the car along the highway, probably because of concern about their safety. Of course, all these inferences lead to conclusions that are merely hypotheses or assumptions. Any of the assumptions made could be wrong, and the inferences by which they were reached are defeasible. So the question is how the inferences are drawn by some kind of logical reasoning.

Early formal systems for planning used formal deductive inferences similar to the models of deductive reasoning used in theorem-proving. Planning was seen as a sequence of reasoning from an initial state to a goal state. The goal state was represented as a set of propositions. The initial state was represented as another set of propositions describing a given situation. Transformation formulas were then used to describe how actions in a plan transform the given situation through a sequence of actions that move towards the goal state. The operation of planning was described as the process of searching for the sequence of transformations starting from the initial state that will end in the goal state. The classic example is blocks world. In blocks world the initial situation is a pile of numbered blocks on a table. The goal state is to put them in several stacks, so that the blocks in each stack are in a certain order. For example, the goal state might be one in which all blocks from one to four are in ascending numerical order in one stack, and all remaining blocks are in descending numerical order another stack.

The problem of plan recognition arises in the kind of case in which an observer of the agent who is performing actions in a blocks world tries to understand what plan this agent is working on. In many typical cases in blocks world, the problem of plan recognition could be easy to solve. One might assume that the final stack of blocks was the goal state, and it might be easy to see how each action

taken led to that goal state. But there is the possibility of error. An agent, even in blocks world, might have had a different goal state in its plan, or might even be trying to deceive the observer. In any real case, like Carberry's example of the car on the highway, there are all kinds of possibilities of drawing an erroneous conclusion. So then we are led to a different problem of plan recognition. How can plan recognition inferences drawn by an observer agent be verified or falsified by evidence? The problem is that the observer can't get into the primary agent's mind, and see what he is planning directly. If these inferences are not tested by observation then, what sort of evidence should be used to evaluate them? The answer given by Carberry (1990) was to build a system called TRACK that uses a framework of dialog to verify or falsify inferences about plan recognition. According to Carberry (1990, p. 75), "TRACK assimilates utterances from an ongoing dialog and incrementally updates and expands the system's beliefs about the underlying task-related plan motivating the information-seeker's queries." In other words, TRACK works as a method of testing and evaluating the defeasible inferences drawn in plan recognition by means of a dialog between the two agents. The plan recognizer has a dialog with the planning agent, asking him questions about his goals and actions, based on what was observed. The plan recognizer does not have to take the planning agent's answers at face value. Various other factors, like apparent inconsistencies, or obstacles to the carrying out of a goal in the plan, can be taken into account. In other words, current plan recognition technology is based on an underlying dialog structure.

What sort of dialog seems to be involved in planning recognition technology? According to Carberry (1990, p. 3), it is an information-seeking type of dialog in which one participant is "seeking information" and the other is "attempting to provide that information". But there is another aspect to plan recognition dialog. Typically, the one agent is trying to get information from the other agent in relation to some actions or tasks that the other agent has carried out or is carrying out. For example, in apprentice-expert dialogs, the apprentice may perform part of the task, and then during an interval, discuss with the expert how to carry out the rest (Carberry, 1990, p. 4). Thus it is apparent that two types of dialog are involved. The primary one is a deliberation, in which the primary agent is using reasoning to carry out a plan. The secondary one is an information-seeking type of dialog in which the plan recognizer tries to make sense of the sequence of reasoning that took place during the deliberations. The deliberation dialog is vitally important, because the plan recognizer must have the capability of grasping how sequences of actions make sense as routine ways of carrying out a goal. The plan recognizer must be able to ask the right questions and to fill in the gaps in a sequence of reasoning. So now we are led to the question of whether dialog theory has the capability of being any use for plan recognition unless deliberation has a structure as a

system of dialog that has clear rules. Fortunately, some recent interdisciplinary research, as noted above, has developed a formal system of deliberation dialog.

Hitchcock, McBurney and Parsons (2001) set out a formal model of deliberation dialog in which two participants engage in argumentation by making proposals and counter-proposals to solve a problem. As noted above, in their model, a deliberation dialog arises out of a need to take action, as expressed in the form of a governing question like "How should we respond to the prospect of global warming?" Both participants in the dialog are agents who have goals and who are presumed to have some capability to take actions designed to achieve these goals. The following sequence quoted from (Hitchcock, McBurnery and Parsons, p. 7) gives the reader an idea of what the main moves in a formal deliberation dialog are, and how each of them is used to contribute to the progress of the dialog.

> Open: Opening of the deliberation dialog, and the raising of a governing question about what is to be done.
>
> Inform: Discussion of: (a) the governing question; (b) desirable goals; (c) any constraints on the possible actions which may be considered; (d) perspectives by which proposals may be evaluated; and (e) any premises (facts) relevant to this evaluation.
>
> Propose: Suggesting of possible action-options appropriate to the governing question.
>
> Consider: Commenting on proposals from various perspectives.
>
> Revise: Revising of: (a) the governing question, (b) goals, (c) constraints, (d) perspectives, and/or (e) action-options in the light of the comments presented; and the undertaking of any information-gathering or fact-checking required for resolution. (Note that other types of dialogs, such as information seeking or persuasion, may be embedded in the deliberation dialog at this stage.)
>
> Recommend: Recommending an option for action, and acceptance or non-acceptance of this recommendation by each participant.
>
> Confirm: Confirming acceptance of a recommended option by each participant. We have assumed that all participants must confirm their acceptance of a recommended option for normal termination.
>
> Close: Closing of the deliberation dialog.

In their formal model of deliberation, unanimity of the participants is required for a decision on a course of action at the end of the dialog. There are also rules for the various locutions (speech acts) that can be made at each move, and rules governing the order of such moves.

But it is not just information-seeking and deliberation that are required for plan recognition. For what happens in a collaborative deliberation between two agents on what to do in a situation where they disagree on who should carry out which necessary task? They may need to negotiate. Or what happens if they disagree about the facts of the situation? To resolve a conflict of opinions of this sort, it can be extremely useful for the agents to have a critical discussion. Just these sorts of conflicts can arise in multi-agent planning. An example given by Chu-Carroll and Carberry (1995, p. 111) shows how negotiation and persuasion dialog often need to be involved as well as deliberation and information-seeking dialog. In this example, the air-traffic control systems in a country fail, and two neighbouring countries, X and Y, negotiate to track and deal with all affected flights. In this scenario, the agents (computerized air-traffic control systems) must engage in collaborative planning together, in order to solve the problem. They are basically involved in deliberation dialog with each other to try to find the right course of action to solve the problem of temporarily restoring air-traffic control to all affected flights. But of course, in order to achieve this goal, they will have to exchange crucial information, for example about which terminal is most capable of guiding which flights. But it is not just information-seeking and deliberation types of dialog that are involved. The air-traffic controllers in countries X and Y will also have to negotiate with each other. If one thinks it is taking on too many flights during a certain time frame, it will have to try to get the other, for example, to trade off by taking some of these flights for others during less busy times. Thus Chu-Carroll and Carberry proposed a model of collaborative negotiation for conflict resolution to provide a multi-agent structure for dealing with agent communication in this kind of situation. But what is perhaps less obvious is that persuasion dialog will have to be involved as well. For suppose the controllers for X and Y have a disagreement in their beliefs about the situation. How are they going to resolve such a conflict of opinions so that their deliberations on how to best solve the problem can continue? The answer, of course, is that they will have to try to engage in reasoned persuasion dialog with each other. Chu-Carrol and Carberry acknowledged this point (1995, p. 121) by recognizing that there are intervals in communication where one agent needs to justify beliefs to another agent. The one agent must use the accepted beliefs of the other agent to carry out an appropriate justification to convince the other agent to accept a new belief. Thus even a basic example of multi-agent planning like that chosen by Chu-Carroll and Carberry shows that at least four of the six types of dialog classified above are necessary to implement multi-agent planning.

7   The BDI model and the commitment model

One general approach to formulating conversational policies is called the BDI (belief-desire-intention) model. The BDI model is based on the central notion of an agent that reasons towards achieving its intentions based on incoming perceptions that update its beliefs. According to the leading exponents of the BDI model (Bratman, 1987; Bratman, Israel and Pollack, 1988; Wooldridge, 2002; Paglieri and Castelfranchi, 2005), an agent possesses a set of beliefs that are continually being updated by sensory input coming in from its environment, and a set of desires (wants) that are then evaluated (by desirability and achievability) to form intentions. On the BDI model, the agent's intentions are defined as persistent goals that are stable over time and are not easily given up.

Typically the conversational policies in ACL's (agent communication languages) have been expressed in terms of the beliefs and intentions of such an agent. An example of a BDI-based conversational policy from the language Arcol, developed by France Telecom in the early 1990's is given by Singh (1994, p. 41): "In Arcol, agent Avi can tell agent Bob something only if Avi believes it also and can establish that Bob does not believe it." The intent of such a conversational policy is to try to ensure that agents communicate collaboratively and sincerely. It is evident that conversational policies of this sort are needed if agents are to communicate in a way that is useful for computing purposes. But the exact rationale of such policies is obscure. The Arcol policy seems restrictive and somewhat arbitrary. Singh (1998, p. 40) has expressed doubts about the usefulness of such a policy. He has questioned whether it is a mistake to emphasize mental agency (the BDI model) so heavily, because it presumes unrealistically that agents "can read each other's minds". It seems that there could be different conversational policies depending on different kinds of conversations that agents might engage in when they communicate with each other. What appears to be lacking is some kind of overarching structure of different kinds of goal-directed conversations in which agents could communicate with each other for some communicative purpose.

As noted above, the dominant approach to providing a structure in which to embed conversation policies in ACL's has so far been the BDI model. This model has also been the paradigm in analytical philosophy of cognitive science. But as shown in the previous chapters of this book, argumentation theory has adopted a different model. This model, the commitment model, is especially prominent in four sources that have been instrumental in mapping out the central dialectical structure on which argumentation theory is based: (Hamblin, 1970), (van Emeren and Grootendorst, 1992), (Walton and Krabbe, 1995) and (Dascal, 2003). Describing Dascal's approach to commitment in dialog, Greco (2005, p. 219) wrote, "The speaker's commitments do not depend on his/her mental states, but are directly

implied by the speech acts that he/she performs". For example, making a promise commits the speaker to fulfilling it, whether the speaker sincerely has an intention to do so or not (p. 219). On this view, for a speaker in a dialog to change her commitment is not to change her belief, or her disposition about the other party in the dialog. It is to go back in the dialog to where something was previously said by that speaker, and change that.

As indicated in chapter 3, the term 'commitment' refers to something different from belief. Belief is a psychological state. Commitment is different. You can be committed to a statement without actually believing it is true. You are committed to a proposition (statement) when you have gone "on record" by asserting it. But there are all kinds of reasons why someone might be committed to something he or she does not actually believe. She may lie, and just have said something that she does not really believe. Commitment is a very useful tool for studying fallacies, and generally for the analysis and evaluation of argumentation in a given text of discourse. The arguer may not be present to be questioned or examined, and we may never know what his inner beliefs really are. Thus it is much more useful to simply examine the given text of discourse containing his argumentation, and then judge what that commits him to maintaining or defending, in light of what he said or wrote. Commitment, or acceptance as it might be called, is an artifact. But what needs to be recognized is that it is a different sort of artifact from belief.

Hamblin (1970) saw what he called a commitment set (store) as a set of statements like those written on a blackboard. As the dialog proceeds, propositions (statements) are added to this set, or can be taken away from this set, depending on a what participants does at any particular move. If the proponent asserts a particular statement at a given move, then that statement is automatically inserted into her commitment store. A participant can also retract commitment to a particular statement, and then it is deleted from her commitment set. But Hamblin (1970, p. 257) made it very clear that commitment, in the sense of the term appropriate for formal dialectic, is not the same as belief. A commitment, on Hamblin's view, is a statement you have accepted during the course of a dialog, in light of the moves (like assertions) you made during the course of that dialog. In many cases commitment and belief will in fact coincide, but you do not necessarily have to believe that a statement is true, in order to be committed to it. Belief can be taken to refer to rational belief, and therefore can be taken in a way that is partly normative. But even on this view of belief, belief is still at least partly an internal psychological notion. Beliefs, desires and intentions are so-called "mental states". Commitment is a normative notion, defined in a dialectical framework. Commitment is determined by the moves made in a dialog. Each move has certain rules, depending on what type of dialog the move is part of. When you make a certain type of move, like making an assertion, asking a question, or putting forward an argu-

ment, the structure of the dialog determines how your commitment set is modified as a result of your having made that move. Thus commitment is, in an important sense, a normative concept. Once you make a certain type of move in dialog, you are then bound or committed to certain statements, in light of your having made that move. What is vitally important to recognize then is that there are two schools of thought on how argumentation should be evaluated as correct or faulty reasoning. The model of reasoning utilized in argumentation theory is commitment-based. The dominant model of reasoning in traditional analytical philosophy, and in much of the recent work in artificial intelligence and cognitive science, is the BDI model.

8   The problem of retraction

In designing any formal model of dialog, the question of how much latitude a participant should have in retracting his prior commitments is a central decision. In some cases, an easy retraction should not be allowed. For example, if Bob has promised to help Wilma move the furniture out of her apartment, but then phones her the night before the move to retract his commitment, he needs to have a pretty good reason for the retraction. Or if a scientist announces a new discovery, but then retracts it, such a move is not really regarded as acceptable. In other cases, retraction clearly needs to be allowed. For example, if Wilma shows that Ed has now asserted a proposition that contradicts an earlier assertion he made. The form of this kind of dialog is illustrated in Table 1.1 above. In such a case, the proponent needs to retract the one proposition or the other, and should freely be allowed to do so. She should even be encouraged to do so. If the respondent shows that the proponent is committed to a thesis that cannot be defended by rational argumentation, then she should retract her commitment to that proposition, or at least be allowed to do so without penalty. Given both kinds of cases, the problem is to set reasonable restrictions on retraction. Retraction should not be allowed too lightly, but on the other hand it should not be fixed in place too rigidly.

Ruling on retraction of commitment is something that should vary with the type of dialog. For example, in deliberation, a fair degree of latitude in retracting commitment is necessary. The reason is that it is desirable for an agent to have a fair amount of prudential flexibility in deliberation. The reason is that reasoning in deliberation is defeasible as new information comes in. The circumstances of the agent in a deliberation tend to be constantly changing, and the agent must be alert to these changing circumstances and react accordingly. On the other hand, in the inquiry, the ideal is to have cumulative argumentation. Once a conclusion is verified as part of the inquiry, it is supposed to be proved to be true, so that there will

be no need to go back and retract it, as the inquiry proceeds. The aim of the inquiry is to avoid retraction of previous commitments, insofar as it is possible.

Now let's consider the persuasion type of dialog. In this type of dialog, retraction should be allowed fairly freely, but not in all instances. Indeed, the main problem with the systems devised in Hamblin (1970; 1971) is that there are no restrictions on retraction. A participant can reply using the "no commitment" move whenever he or she wants to. The problem is that it is impossible to pin down commitment when engaging in a critical discussion. As soon as you appear to be getting close to disproving the opponent's thesis, or any proposition near to that thesis which could pose a threat to it, he can simply reply, "No commitment." This freedom is a serious problem in a persuasion dialog, because you can never use any premises as "fixed points" to prove your thesis, or to disprove your opponent's thesis. As soon as you try to deploy your argument, the other party will simply retract commitment to the premises. How then can the problem of retraction be solved for persuasion dialog?

There can be many ways of trying to solve this problem. There could be penalties for retraction, for example, that give some advantage to the other party, making his task of persuasion more difficult or lengthy. The solution advocated in (Walton and Krabbe, 1995, pp. 144–149) is to require *stability* in the retraction of commitments, meaning that a commitment cannot be retracted in isolation from related commitments in an argument. How this solution works in practice can be illustrated by a cases of what are called stability adjustments. An *internal stability adjustment* requires that once certain supporting premises for a commitment are removed, the commitment itself may have to be removed, unless other supports can be found. For example, consider the following dialog.

> Bob: The track is not open on Friday.
>
> Edna: How do you know?
>
> Bob: The notice on the board says so.
>
> Edna: I saw the gym monitor remove that notice.

It is appropriate at this point in the dialog for Bob to retract his commitment to his earlier claim that the gym is not open on Friday. The reason is that his reason for that commitment is undercut or at least challenged by Edna's remark. Since the monitor presumably knows when the gym is open or not, his action of removing the sign is quite good evidence on whether the gym is open or not. In this case then, an internal stability adjustment requires that Bob should retract his commitment. On the other hand, suppose the dialog were to continue as follows.

> Bob: Yes, but I saw him put it back up again.

In this case, Bob has given a good reason why he should not have to retract his commitment. So we see the general problem quite clearly here. Commitments are not always subject to retraction, but they are also not always non-retractable (fixed). The best rule for retraction lies somewhere in between.

An *external stability adjustment* (Walton and Krabbe, 1995, p. 147) requires that once a particular proposition is retracted, some premises or warrants leading into that commitment will also have to be retracted. How such external stability works can be illustrated by the following simple case.

Consider the following sequence of dialog that occurs within a critical discussion on the issue of tipping. Bruce argues for the thesis that tipping is generally a good thing. Wilma thinks that tipping is a bad thing. As the argument proceeds, it becomes clear that she even thinks that tipping is a social practice that ought to be abolished. Bruce questions this point.

> Bruce: Why should tipping be abolished?
>
> Wilma: If something has bad consequences, it should be abolished. Tipping has bad consequences.
>
> Bruce: Well that is a valid argument. But why does tipping have bad consequences?
>
> Wilma: Tipping leads to many misunderstandings. Anything that leads to misunderstandings has bad consequences. Therefore tipping has bad consequences.
>
> Bruce: Well, OK. Once again, your argument appears to be valid.

But then, let's suppose, the dialog goes on for quite an interval, and Bruce produces many convincing arguments to the effect that abolishing tipping is not really feasible. Wilma then gives in. Instead of trying to argue that tipping should be abolished, she takes the line that it ought to be regulated and standardized. Although she thinks that tipping is bad thing, in general, she concedes that we can never entirely get rid of it, human nature being what it is. So she takes the line that it ought to be reconfigured in a form that eliminates or minimizes the bad consequences.

Now the problem in this case is – what should be done about Wilma's commitments, as expressed in the dialog sequence above? Should she just be allowed to make the retraction, and should that be the end of it? Or should there be some penalty attached? The method of external stability adjustment would require a kind of penalty, in the sense that Wilma would not only have to retract this single proposition from her commitment set. She would also have to retract certain other propositions closely related to it. To see how the method works, it is best to construct an argument diagram.

Wilma used a chaining of two deductively valid arguments, as represented in the argument diagram below (figure 1.2). The five propositions in the argumentation are labeled as follows.

A: Tipping should be abolished.
B: If something has bad consequences, it should be abolished.
C: Tipping has bad consequences.
D: Anything that leads to misunderstandings has bad consequences.
E: Tipping leads to misunderstandings.

The sequence of argumentation is represented in figure 1.2.
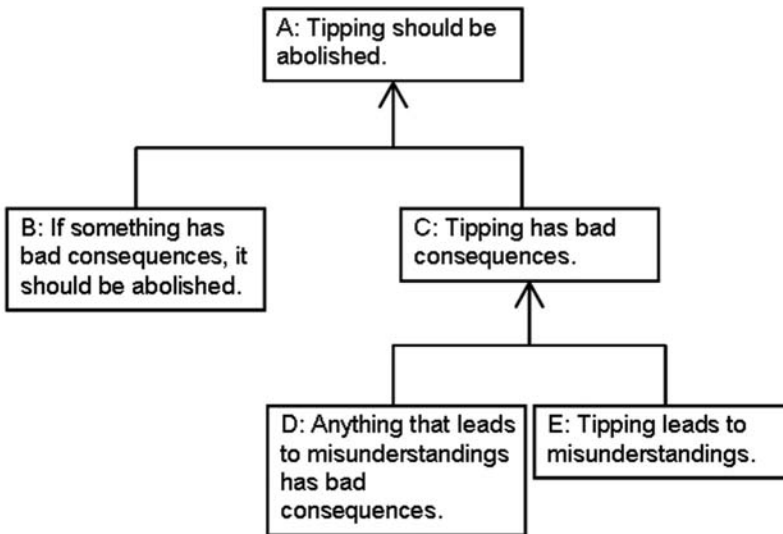
Figure 1.2 Argument diagram of the tipping argument[17]

Now that Wilma has retracted commitment to *A*, what should be done about her commitment to the other propositions represented in the diagram? *B* and *C* together, as premises, deductively imply A. If A is false, then one or the other of *B* or *C* must not be true. It appears then that Wilma should be given a choice. If she retracts commitment to *A*, then she should either retract commitment to *B* or retract commitment to *C*. What if Wilma decides, after critical questioning by Bruce,

---

17. The diagram in figure 1.2 was drawn with *Araucaria*, an automated system of argument diagramming explained in chapter 7, section 4.

to retract commitment to *C*? Then she should face another choice. She must now either retract commitment to *D* or retract commitment to *E*.

The general policy of stability adjustment as a solution to the problem of retraction is as follows. If an arguer in a persuasion dialog retracts commitment to a conclusion of a linked argument, then she must retract commitment to at least one of the premises. If the argument is convergent, she must retract both (or all) premises. If the argument is serial, then she must go back up the chain of arguments, as indicated in the example above, and make all the required retractions along the way, as far as the chain extends. This policy on retractions does impose a kind of penalty on the arguer, so that she will not make retractions of commitments too freely. For unless she also retracts the related propositions in her commitment store – which she may not want to do, and which may take some time and effort – she is not allowed to retract any particular proposition. The general policy means that an arguer is free to retract commitment to any propositions that are not intimately related to his central position, but is more restricted when it comes to retraction of propositions that are more centrally related to his main lines of argumentation within his commitment set.

At any rate, stability adjustment is one proposed solution to the problem of retraction in persuasion dialog. Whether it will turn out to be the best solution remains to be seen. Whether there may be different kinds of persuasion dialog, in which different solutions for the problem of retraction are more useful, also remains to be seen. For the moment, the problem of retraction has been posed, and one proposed solution has been shown. Future work on dialog theory will be concerned with devising appropriate rules of retraction for the various types of dialog represented in Table 1.2 above.

9   Communication and information

Dialog theory looks to be extremely useful to model the fundamental notion of communication that is the basis of so much human activity, in business, advertising, sociology, and so many affairs of life and learning. But communication, judging from what has been written about it in the field of communication studies, has been a problematic notion to define. Communication has been defined in many different ways by the many authors who have written on the subject of communication (Craig, 1999), but central to many of these definitions is the notion that communication is the transfer of information. For example, according to the definition of 'communication' presented by Kimura (1993, p. 3), "The term is used here in a narrower sense, to refer to the behaviors by which one member of a species conveys information to another member of the species." But what exactly is

information? According to the mathematical definition given by Shannon and Weaver (1972), information is defined as the reduction of uncertainty, measured by the change in probability values in a field of events. On this definition, information is independent of human perception and interpretation. It can be measured quantitatively as bits being transferred over a wire from one terminal to another. What this mathematical formula is supposed to measure is whether the type of response selected by a receiver is causally influenced by the type of message selected by the sender (Hauser, 1996, p. 8). Whether this narrow and technical concept of information is adequate to represent the right meaning of the term needed to understand the structure of information-seeking dialog is an open question.

The more important observation here is that from the viewpoint of dialog theory, information-seeking dialog is just one type of dialog framework of communication in which argumentation takes place. The persuasion dialog, the inquiry, deliberation, negotiation and eristic dialog also represent important kinds of goal-directed communication. But surely the goal of each of these five types of dialog is not just the conveying of information. Well, yes, that is right, according to the new dialectic, but of course it all depends on how you define 'information'. If that term is defined broadly and inclusively enough, and perhaps also neutrally enough – for example, in terms of increase or decrease in probability values in a series of messages – then just about anything that happens in a dialog could count as a transfer of information.

Let's take the case of persuasion dialog, because that is a type of communication that has been seen as especially important for rhetoric as a subject. The goal of the dialog as a whole is to resolve a designated conflict of opinions, or at least to reveal the strongest arguments on both sides of the issue. As indicated above, the goal of the one party is to get the other party to accept his or her (the first party's) thesis, which is one of the propositions on the one side of the conflict of opinions. This one party tries to accomplish this task by using arguments that have commitments of the other party as premises. The task is carried out in a series of small steps making up a longer, connected sequence of moves. That is the essence of persuasion dialog. But now how can persuasion dialog, so conceived, be seen as a form of communication that represents a transmission of information from the one party to the other? Well, one could say that persuasion, so conceived is a transmission of information, because each party is coming to find out all the reasons that support not only her own thesis, but also that of the other side. And this finding of reasons, or supporting arguments, is a kind of increase of information. This move seems fairly plausible (not to me, but it will to many readers). So here the problem is that it is hard to exclude any kind of communication from being classified as a transfer of information, provided the term 'information' is construed broadly and inclusively enough. The problem is exactly that. The term 'informa-

tion' is currently, in fact, used as such a flexible and catch-all category that it hardly seems to exclude anything encompassed by any form of communication that has a cognitive aspect. For example, if you put the key word 'information' into a search in a library catalogue or other bibliographical data base, you will get huge numbers of hits with that word in the title. Most of them will be in the field of computing, and most of them will have little if anything directly to do with the narrower sense of the term used in the new dialectic. 'Information' is one of those terms like 'relevance' that has become a so-called "rag-bag" category into which you can conveniently throw anything that is unclear or not well understood.

The solution or way out of this problem is to overcome the old positivistic idea that information is something solid and objective, like "the facts", that needs to be measured objectively, quantifiably and operationally, as bits transferred over a wire, or as increase or decrease in probability values. Instead, information-seeking needs to be seen as a distinctive type of dialog concept in which a set of propositions is transferred from one party to another. But it is not just any old set of propositions that should qualify. Information needs to be seen as a set of propositions that supposedly represents the real facts of a case, or at least part of a plausible account of what happened in a case. The concept is simple enough in some cases. For example, if I ask you what color Cher's hair is, you may provide that information by stating the proposition, 'Cher's hair is black'. If Cher's hair is in fact black, then you have given me the requested information. As shown in chapter 5, exchange of information in a simple kind of case like this can be managed with an "ask-tell" conversational move (speech act). But this simple question-reply structure cannot effectively deal with more complex kinds of question-reply exchanges that are very common and useful in communication. The problem with many typical cases is that I may need information, but I am not sure exactly which specific propositions I need to find, or what questions I should ask, right at the beginning of the search. For example, if am starting a research project on air pollution in Chicago, I am not sure what I might find. So I just start searching around for information that seems relevant. When I get relevant information, I then try to evaluate its accuracy and reliability by comparing it with other information, and otherwise critically judging its worth. Thus the simple "ask-tell" unit of communication is not adequate to this more sophisticated method of collecting information. Looking at communication on the model of the simple transfer of propositions that are true or false is too narrow.

Once the concept of communication is viewed in a broader way than this narrowly positivistic way, it becomes clear how information-seeking dialog is an important type of dialog in its own right, as a framework of argumentation. There are other highly significant types of dialog as well that need to be considered under dialog theory as representing distinctive types of communication. But once the

problem of modeling information-seeking dialog as one type within other types of dialog is solved, the way is clear to adopt dialog theory as the core structure underlying not only argumentation theory but communication theory as well. Thus now dialog theory is coming to be accepted as a formal structure in which argumentation can be modeled, it is getting a life of its own as a theory. Once dialog theory is more widely accepted, it will become quite natural to see it as the most natural and useful model of communication generally.

10   The future and past of dialog theory

It is possible to see from the range of subjects covered above that dialog theory will have to be a broad subject that encompasses argumentation in different types of dialog. Six dialog skills have been shown to be especially important: (1) how to enter into a dialog, (2) how to recognize a type of dialog, (3) how to productively engage in a dialog with a collaborative speech partner, (4) how to avoid being deceived by a speech partner who only pretends to be collaborative, (5) how to be flexible in shifting from one type of dialog to another, (6) how to close off a dialog. Consider an agent that has to act with other agents in an open and complex society where there is an abundance of information, and where carrying out a task requires collecting that information by engaging in dialog with other agents. Sometimes, for example in academic research, an agent will do best by trusting another agent to be cooperative and to tell the truth. But in other contexts, for example in commercial negotiations, it may be more productive to assume that another agent will act in a self-interested way, even if it involves not telling the truth, or not paying much attention to the truth. Teamwork requires one kind of commitment based on trust and reputation, but avoiding fraud and deception requires a critical vigilance. The agent's argumentation skills must include the ability to detect contradictions and gaps in the arguments put forward by another agent, and to draw the right implications, depending on how the other agent responds when questioned about these apparent problems. Thus an intelligent agent will open a dialog and stick to the norms and the forms of argument appropriate for this type of dialog. But at the same time, the intelligent agent will be ready to close off the discussion, or shift to a different type of dialog, if that is what the situation calls for.

Dialog theory thus opens up a whole new range of problems that have not been really considered before in any systematic way. There are all kind of special problems relating not only to each type of dialog, but also to entering into a dialog, and how to move from one type of dialog to another. At a more abstract level, taking the new subject as a whole, is the very idea of dialog itself as a form of communication and as a framework of rational argumentation. This new structure af-

fects fields as widely varied as computer science, communications, logic, artificial intelligence, legal argumentation, cognitive science, sociology, linguistics, and anthropology. The foundational notion is that of the dialog, and the idea that an argument can be judged as rational or not on the basis of how well it contributes to a dialog that the arguers are supposed to be engaged in. Instead of the abstract notion of rational argument as knowledge or belief justified by context-free inference, we move to a social notion of rational argument in dialog. An argument is then rated on the basis of how well it contributes to a goal-directed dialog of some identifiable type in which two speech partners are taking part. Feminists and others who have seen traditional logic as too abstracted from a social context, too rigid, and less tuned to real social practices of collaborative communication and teamwork, will welcome the advent of dialog theory as a liberating influence.

Looking over the various matters covered in chapter 1, one can see that dialog theory shows a lot of promise, as a kind of structure that could applied to many areas and problems. Formal dialectic as applied to different types of dialog offers quite a general theory of a kind that has been lacking in the past. It shows special promise as applied to multi-agent computing and communications. On the other hand, dialectic seems to be such a new field that it may be hard to take it seriously. Especially for those of a logical empiricist way of thinking, the whole idea of relativizing rational argumentation to different types of dialog may sound suspiciously subjective, even postmodern, to use that philosophically controversial term. Surely, such critics might maintain, taking such a dialog relative view of argumentation is running the risk of blurring the line between rhetorical argument and persuasion on the one hand, and objective truth and valid argument on the other hand. Shouldn't we be wary of any supposed new field that judges the rationality of an argument in light of how it is used for purposes of persuasion of an audience? Audiences are notoriously prone to being persuaded by arguments that are emotionally compelling but fallacious. The blurring of the line between rational argument and sophistry could be a result of moving to the dialectical notions of commitment or acceptance, and moving away from the more respectable notions of truth and validity.

Answers to these worrisome doubts can be given but they are not trivial. Giving reasoned and adequate answers means going back in history to the birth of logic and rhetoric as systematic fields springing from Greek philosophy. It is well known that rhetoric is an old field that was first systematically studied in a probing way by Aristotle. What is less well known is that Aristotle also built on the foundations provided by his predecessors to claim the existence of an important field called dialectic, based on a dialectical argument. As noted above, Aristotle also had sophisticated views about the relationship of rhetoric to dialectic. So dialog theory, or dialectic as it was known to the Greeks, is actually not an altogether new field.

Thus in order to rehabilitate dialectic, or to gain it some respect as a potentially significant field of academic study and research, it is necessary to go back to the Greek philosophers. It is necessary to grapple with some ancient questions concerning the relationship of dialectic to rhetoric. This historical investigation of the conceptual roots of both subjects is undertaken in chapters 2 and 3. In chapter 2, the history of dialectic as a subject is traced from its Greek roots through the middle ages to modern views of it. Even to see that the Greeks founded the subject and thought it to be important is revealing. Because of the convoluted history of western thinking about logic, the notion of rhetorical argument has been slanted in a certain way, reinforcing the old idea of Plato's that rhetoric is the enemy of philosophy. It will be argued in chapter 3 that this old but pervasive view is based on a false opposition that needs to be overcome. It needs to be seen that rhetoric is a necessary part of dialectic and that dialectic can also be an extremely useful part of rhetoric. It takes pains to develop this new approach to rhetoric and dialectic, and overcome a lot of historical antipathy and mistrust. But once it is done, a new way of looking at both subjects becomes possible. There is revealed a new way that makes rhetoric a much more powerful subject, based on argumentation structures that give it a central methodology. And there is revealed a way of reviving Greek dialectic that makes it a powerful new form of applied logic that can be applied to the interpretation and analysis of argumentation in natural language discourse.

The history of dialectic

Dialectic is a very old subject (or technique or art) that was a central method in ancient Greek philosophy. Different Greek philosophers defined it in different ways, but the common element is the notion of two (in the simplest case) parties reasoning with each other. Perhaps dialectic could broadly be defined, after the Greek conception of it, as the art of rational argument by conversation. Dialectic began to be developed as a technical part of logic as the Obligation Game in the Middle Ages, but stayed more or less to the Greek conception of it, and did not lead to any influential method. Dialectic then failed to survive, in any robust way in its Greek form, in modern philosophy. Kant took dialectic to be the critique of illusion, and thought it to be not much different from sophistry. Hegel took dialectic in a different direction, as concerned not just with verbal exchanges but with the way things in themselves really are. In the twentieth century, dialectic became an obscure subject that had to be described mysteriously to the unfamiliar as a kind of ancient "mental gymnastic" for argumentation training. But now, with the advent of recent developments in argumentation theory, artificial intelligence, and linguistics, dialectic is making a comeback. In modern terms, it is seen as a branch of applied logic that has to do with the evaluation of argumentation in various contexts of conversational use. Its special use is the analysis and evaluation of fallacies.

Now several questions are posed. What is dialectic? Or what should we take it to be, as a well-defined subject that could potentially be of some use? Is the modern version of it similar to the Greek version, or the various Greek versions, from what we know of them? Why did dialectic fade out in the modern world, and become obscure? Did dialectic take a bad turn, with Kant and Hegel, that made its Greek roots even more remote? Would a "Greek revival" of dialectic be appropriate? Or is the modern version we require now different altogether in quality from the ancient Greek idea of it? Is there any common element to the various formal systems of formal dialectic? If so, can this common element be used as a basis for defining dialectic as a subject or method? These are the questions considered in chapter 2. The question of how dialectic should be compared or contrasted with rhetoric is taken up in chapter 3.

The question of how dialectic should be defined, as some activity or some co-herent field of study, is timely now that many systems of formal dialectic have been constructed in logic. The proliferation of different formal systems calls for some central conception of dialectic that would be appropriate for modern develop-ments. It would be simplistic to think that the new dialectic necessarily has to be the same thing as the old dialectic, as defined in the ancient world. And yet it is quite clear that dialectic very definitely has ancient roots. What then is the rela-tionship between the old and new dialectic? And what meaning should be as-signed to dialectic as a serious method or logical structure that could be of any use or scholarly or scientific interest for modern logic? These questions are only fully answered later, as other chapters deepen our knowledge of both the old and the new dialectic. But to have a good point of departure, a central concept of what dialectic is, or should be, needs to be developed.

1   Origins of dialectic in ancient philosophy

Dialectic was taken by the Greeks to be a branch of what is nowadays called logic. Since that time, however, logic has not centrally been thought to be about how two persons reason in a conversation when arguing with each other. But dialectic was taken by the leading Greek philosophers to be a kind of applied conversational logic of just this sort, and they took it to be extremely powerful and important.[1] Socrates called dialectical discussion the greatest human good (*Apology* 38a). The Stoics thought that the wise man must be skilled in dialectic. Diogenes Laertius, describing the Stoic view of dialectic, wrote: "Without the study of dialectic the wise man will not be infallible in argument since dialectic distinguishes the true from the false, and clarifies plausibilities and ambiguous statements." (Long and Sedley, 1987, p. 184). Dialectic appears to have been an important part of most of the ancient philosophies, but there was a particular school of philosophy called the

---

1.   The two key Greek terms associated with dialectic are *dialegesthai*, translated by Timmer-man (1993, p. 117) as "the activity of holding a discussion", and *dialektike*, translated by Tim-merman (1993, p. 117) as "the art of dialog". The term *dialegesthai*, which seems to be the root of *dialektike*, came to be used in a technical sense in Greek philosophy, but it also has a more ordinary meaning. In the *Apology* and other dialogs, Socrates often refers to his conversations by using the word *dialegesthai* "with no suggestion that the word carries any philosophical weight" (Kahn, 1998, p. 303). But in other dialogs, like the *Gorgias*, *Protagoras* and *Hippias Minor*, there is a contrasting "methodologically marked" use of *dialegesthai* by Socrates to refer to the techni-que of question and answer, in contrast to rhetoric or oratory as used by the Sophist (Kahn, 1998, p. 303). The adverb *dialekitkos* refers to someone who is skilled in philosophical conversa-tion, and could be translated as "dialectician" (Kahn, 1998, p. 303).

dialectical school. It was active around 350–250 BC, and was an important precursor of Stoic logic (Sedley, 1998). This school stressed "the intrinsic value of dialectical activity – argument by question and answer" (Sedley, 1998). Most often dialectic is associated with the question-answer technique of argumentation employed by Socrates in the Platonic dialogs.

Dialectic (*dialektike*) was Plato's name for the kind of systematic discussion by question and answer (*dialegesthai*) that is seen to be practiced in the Socratic dialogs (Irwin, 1988, p. 7). Socrates questions various interlocutors in the dialogs, typically on matters concerning ethical opinions or other controversial philosophical issues, and the discussion turns up logical problems, often leading to puzzlement. The Socratic dialog provides a paradigm for Greek dialectic. But in the ancient world, opinions differed on exactly how to define dialectic. Even within the development of one philosopher, 'dialectic' may evolve from one meaning to another. According to Robinson (1962, p. 70), the meaning Plato assigned to the term 'dialectic' changed over time, and he tended to see dialectic as "the ideal method, whatever that may be." It would be simplistic to think that there was a unified meaning of the term 'dialectic' amongst the ancient philosophers, or that the old dialectic can be simply defined as a method shared by all the ancient schools. But at least from the early Platonic dialogs, a fair idea of what dialectic was supposed to be, as a philosophical method, is made clear by example. Socrates picks as an interlocutor someone who professes to know the truth of a matter. Then by a process of using a sequence of questions and replies, he tests out the claim to knowledge. This dialectical process is usually called *elenchus*, or refutation, in Greek philosophy.[2] How does it work? How it works (Robinson, 1962) is that Socrates remembers the earlier replies of the interlocutor. As the process of questioning proceeds, he makes use of these earlier commitments by comparing them to the commitments presented by the answer just given. Such comparisons can reveal weaknesses and questionable assumptions. In some cases, Socrates even finds what appear to be contradictions. An inconsistent set of commitments cannot collectively be true. Therefore, the interlocutor in such a case must deal with the apparent contradiction by refining his view. The process of refinement of the initial view, as the dialog proceeds, is the development whereby the dialog leads towards the truth of the matter discussed. Exactly how this progress works is somewhat mysterious. Socrates often talks about recollection (*anamnesis*). But ba-

---

2.   Hence the odd sounding Greek expression meaning fallacy or sophism is *sophisticus elenchus* or "sophistical refutation". The title of Aristotle's book on fallacies is *On Sophistical Refutations*. This ancient terminology seems as mysterious and incomprehensible to the modern reader as the term 'dialectic' itself. But once you get used to the notion of dialectic in the Greek sense, it does tie in well with the notion of fallacies as being described as sophistical *elenchi* or refutations.

sically the process of dialectic seems to work because the participants really care about the truth, and therefore collaborate with each other in the dialog, even though there is an adversarial element as well.

Where did dialectic originate? There are two ancient sources reporting that Aristotle said that Zeno of Elea was the inventor of dialectic. According to (Kneale and Kneale, 1962, p. 7), both Diogenes Laertius and Sextus Empiricus made this report. But this reported remark of Aristotle could be misleading. Much depends on what is meant by 'dialectic'. This term is usually taken to refer to a dialog that takes place between two parties (in the minimal and standard case) in which they take turns asking and answering questions and putting forward arguments. Zeno is famous for having invented the *reductio ad absurdum* technique of argumentation in which a proposition is refuted by showing that it leads by logical reasoning to an absurdity, another proposition that couldn't be true, like a contradiction. Thus Aristotle's remark may only mean that Zeno invented the method of examining a hypothesis by drawing logical consequences from it. This method does have dialectical aspects, because it involves a sequence of logical reasoning in which one party can draw logical consequences from the hypothesis put forward by another party. It is also a method used by Socrates in the Platonic dialogs. He often draws out logical consequences of another party's stated view in a dialog, showing that the line of reasoning ends in an absurdity or difficulty. As noted above, this Socratic method of questioning a respondent in a sequence of dialog that leads from the respondent's answers to a particular conclusion drawn from them as the ultimate outcome of the sequence is called the *elenchus* (from the term *elenchos* or refutation). But the method of *reductio ad absurdum* is surely not all there is to dialectic or to the *elenchus*. *Reductio ad absurdum* is used in Euclidean geometry, for example, in a way that is not necessarily dialectical, in the sense of involving two parties engaged in question-reply dialog. Thus Aristotle's reported remark may be misleading, in certain respects. Robinson (1962, p. 91) agreed that the statement that Zeno was the discoverer of dialectic is "inaccurate and misleading if we take the word in either Plato's sense or Aristotle's." Robinson suggested (p. 91) that it is "extremely improbable either that Zeno ever thought of the word 'dialectic' or that he ever entertained a method similar to Plato's or Aristotle's dialectic." The reason is that although one of Zeno's arguments has reached us in a question-answer form, there is no evidence for the thesis that Zeno thought that dialog by question and answer was necessary to good philosophical method (Robinson, 1962, p. 91). Thus despite these ancient sources, it is not plausible that Zeno invented dialectic in the sense of the word meant by Plato or Aristotle.

It is a more plausible hypothesis that dialectic originated with the Sophists, teachers who taught their students skills of arguing both sides of a case in which there is a conflict of opinions. The Sophistic work called the *Dissoi Logoi* (Robin-

son, 1979) shows these teachings, and they can be called dialectical, because they exhibit typical use of argumentation in which the arguments of one party are opposed to the arguments of another party in a dispute. Some might say, however, that Sophistic manuals like the *Dissoi Logoi* are only rhetorical, rather than dialectical in nature, and that dialectic proper was only discovered by Plato and Aristotle. Another hypothesis is that Socrates may have invented dialectic. Again here much depends on what you take the term to mean.

> There is no reason to doubt that Socrates was a master of the art of philosophical conversation by question and answer, as illustrated in the dialogs. But there is every reason to doubt that Socratic practice had the rule-governed form sometimes attributed to the *elenchus*. (Kahn, 1996, p. 302).

This remark suggests the hypothesis that Plato should really be considered the inventor of dialectic. But even this claim, despite its plausibility, is ambiguous. It could mean that Plato invented the notion of dialectic found in the Platonic dialogs, or it could mean that Plato invented the technical term 'dialectic' as a logical concept (Robinson, 1953, p. 90). Robinson (p. 90) argued that both of these hypotheses are "probably true". On the other hand, it is possible to argue that although Plato gave the best and most lasting examples of the use of dialectic in his dialogs, he didn't define the term in a precise and consistent enough way so that it could useful for logic as a technical discipline. It was left to Aristotle to do that. Perhaps then the attempt to credit a single individual as the inventor of dialectic does not do justice to explaining how the concept came into use in Greek philosophy, and how it came to be developed as a logical tool that can be used in philosophy.

2   The dialectic of Socrates and Plato

Robinson (1962, p. 77) outlined the main characteristics of the kind of conversation Plato held to constitute dialectic as the supreme method of philosophy. These various characteristics can be put in the form of rules. How strict they should be taken to be, as rules, is debatable. Nonetheless stating them as general rules of Platonic dialectic is a useful exercise, because it raises questions about whether, or to what extent, the Platonic model of dialectic can be formalized or codified by clear and precise rules. The first rule (p. 77) is that dialectic must be a social activity involving at least two participants. This statement seems to imply that dialectic cannot be carried out by an individual alone. But of course, this statement always raises the issue of whether a single person could play both roles, that of the proponent and the respondent. Some would cite the evidence of "devil's advocate" discourse, in which someone trying to decide what to do or what to believe lists the

reasons both for and against. In such a case, there is a kind of dialog or dialectic taking place, but there is only one person in it. A way out of this impasse is to distinguish between a participant and a person. The same person could play two roles in a dialog. In effect, there could be two participants, each with a different role, even though there is only one person in the dialog. What is shown, apparently, is that the notion of a participant, often called a "party" in a dialog, is more closely linked to the role or function taken than it is to the actual number of persons. To cite another kind of contentious case, you could have two groups of persons on either side discussing a contentious issue. And yet each group could represent a viewpoint that defines it as a participant or party in the dialog.

The second rule is that the respondent or answerer is "expected to say what he really thought, and nothing else" (p. 78). But this rule is not absolute. It is affected by two other constraints (p. 78). One is that the respondent is expected to be consistent in his answers. Thus there could be conflict between maintaining consistency and saying what you really think is true. This conflict is the source of the Medieval Obligation Game, outlined below. The other constraint is that there is supposed to be agreement between the proponent and the respondent. The third rule (p. 79) is that the respondent must always answer the question, and cannot just plead ignorance. The value of this rule is that it helps the discussion along to probe into the issue, and prevents the respondent from evading the question. But Platonic dialogs apart, this kind of rule is quite controversial, and it is difficult to know how to implement it in a real case. The main problem is that questions can be loaded with presuppositions. Forcing the respondent to give a direct answer to every question in every instance could permit the proponent to commit fallacies. For example, suppose the proponent asks the famous spouse abuse question, "Have you stopped abusing your spouse?" By the Platonic answer rule, presumably the respondent has to answer 'yes' or 'no'. But suppose he has never abused his spouse, in his opinion. Then he is prevented by the Platonic answer rule from giving his real opinion. And so he is forced to violate the second Platonic rule. Thus there are problems concerning whether this rule can be codified or formalized in any precise way that does not turn out to run contrary to the aims of Platonic dialectic. The fourth rule is that "dialectic recognizes no authority." (p. 79). The respondent is not allowed to accept anything on the basis of authority. Instead he must answer on the basis of "what seems true to us here and now" (p. 79). This rule, along with the second rule, tells us that dialectic is supposed to be about the opinions of the respondent.

These four rules narrow down dialectic, but still don't really tell us what it is. To see the aim of dialectic, we have to look at what Socrates does in the dialogs. His aim does seem to be to discuss a particular issue or question that is controversial, and that admits of views that are opposed to each other. But he doesn't just discuss the issue in a random manner. His aim seems to be to try to get to the truth of the

matter. Of course, especially in the earlier dialogs, he tends not to be very successful, or only successful to some extent. He does not come up with a final conclusion that is beyond doubt. He sometimes refutes a particular view being discussed. And even when the discussion is favorable to a view, it supports it more as a provisional hypothesis that seems plausible, than as a conclusion that has been proved beyond doubt. The argument in the discussion supporting the view in question is typically based on definitions that are put forward in the dialog as hypotheses. Socrates gets the participants to agree to a particular definition provisionally, to see where it leads in the dialog. And thus the outcome of the line of argumentation in the dialog depends on that definition as a hypothesis. Of course, Plato thought that definitions are based on essences that are objective, real, fixed and timeless. But especially in the earlier dialogs, the way the definitions are put forward and argued about seems to suggest that they are more provisional and tentative, and that their acceptance is subject to argumentation and agreement. For example, in the *Meno*, no absolute conclusion is reached at the end on what virtue is. And yet the dialog is quite informative philosophically, in that it is quite illuminating on the subject of what virtue is, and is not.

The problem remains then. What is the purpose of a Socratic dialog? Supposedly the purpose is to get to the truth of the matter being discussed. But as indicated above, the dialogs do not appear to do that, or even to purport to do it, by proving something beyond all further doubt. Perhaps then the aim is better seen as one of getting part way to the truth, or closer to it, by refuting the false contenders, the hypotheses that initially seemed plausible but are shown to be false by the discussion. But this aim is negative. The Socratic dialog seems to have more than just a negative aim of refuting false views. How could its positive aim be described or defined? Of course, the dialog is meant to be useful for teaching. But as Robinson (1962, p. 80) observed in connection with the *Phaedrus*, "Plato quite evidently thinks of dialectic as a method of discovery as much as a method of teaching." The question thus can be put as one of how the aim of discovery is supposed to take place. Plato does not seem to say, anywhere in the dialogs themselves, how dialectic is supposed to discover or uncover the truth of a matter being discussed. According to Robinson (p. 81), there is no satisfactory answer to the question of why the question-answer method is essential to discovery in any of Plato's writings.

One clue as to how the question-answer method is part of a discovery process is the observation made by Seeskin (1987, p. 24) that Socrates often refers to *elenchus* as a kind of persuasion. But what is meant by 'persuasion'? It is clear from the four rules of Platonic dialectic cited above that what matters is the opinion of the respondent. By 'persuasion' must be meant the operation of the proponent getting the respondent to change his opinion. In a Platonic dialog, a respondent starts out the discussion with some opinion in the form of a statement that he is committed

to. Then Socrates finds some problem with it, or some questionable aspect of it, and the respondent alters the statement, or perhaps even moves to a different statement as expressing his new view of the matter. So this change from one commitment to a different one is persuasion. But if it is just the respondent's personal opinion that is changed, then how is this process any different from psychological persuasion? And if it is only individual or psychological persuasion that is involved, what value is that kind of change in leading to a discovery that would move toward the truth of the matter? For just because one individual believes some statement or not does not make it true or false. The answer is that the term 'persuasion' here refers to something more than just psychological persuasion. Socrates uses logical reasoning to draw out inferences from a view, and he uses other logical devices, like deriving a contradiction form the view to argue that it is false. Moreover, the views examined by Socrates are not just any personal opinions. They represent plausible opinions on matters that are philosophically controversial. Also, Socrates, in many dialogs, tries to be exhaustive in examining all the plausible or widely known views that might provide an answer to the question being discussed. So while persuasion as meant in Platonic dialectic is partly psychological or individual persuasion, it is also partly a process of rational or logical persuasion. In some sense, persuasion is very important in the structure of a Platonic dialog, and is a key to understanding the purpose of Platonic dialectic. But what is meant exactly by 'persuasion' has been notoriously hard to grasp in the past, and is very controversial. This subject will be more deeply investigated in chapter 3.

Judging from what Socrates does in the dialogs, and from Plato's view of discovery and learning as *anamnesis* (recollection) expressed in the dialogs, it is possible to suggest a plausible hypothesis about the purpose of a Socratic dialog. This hypothesis is not necessarily meant to represent Plato's (or Socrates') view of the matter, but it is a view that is consistent with what happens in the dialogs. This hypothesis is that the aim of the dialog is to get the respondent, through the discussion with the proponent, to refine his own view. Refinement of the view takes place positively by coming to understand the reasons behind it, and negatively by getting the bugs out of it, for example by removing contradictions in earlier formulations of it. The aim is not so much one of proving a view to be true or false (beyond doubt), as one of deepening the respondent's understanding of a view. The aim, so seen, is one of deepening the respondent's insight into a view by probing into the reasons that support it or go against it. The outcome of the dialog is not necessarily that the respondent changes his original view, although that can happen. It is more to be seen in how he gets deeper insight into that view by expressing it in a more sophisticated way that takes into account the leading arguments both for and against it. Thus the respondent will qualify the original view by adding necessary qualifications to it, or by defining the key terms in it in a more precise

and adequate way. Of course, this way of stating the aim of dialectic is simplified in a certain respect. Dialectic is a cooperative activity. It is not just the respondent who achieves greater insight into his view. The proponent's view of the matter is also changed and refined by the process of questioning and answering. Moreover, certain other persons who hear or read the dialog, or even take some part in it, can have their views changed as well.

Accordingly, the concept of dialectic that describes the process that takes place in a philosophical dialog of the Platonic type has what can be called a maieutic function at its center. *Maieutikos* is the Greek word for midwifery.[3] By questioning the respondent, the proponent (Socrates, in the Platonic dialogs), brings new ideas to light that were already there within the respondent to begin with. But by examining these ideas through discussion, critical argumentation and probing questioning, the proponent guides the respondent to improve and refine his initial view. Thus both parties to the dialog have an increase in their depth of understanding of this initial view, and of the most plausible and convincing arguments for and against this view. The purpose of the dialog, so conceived, is to increase maieutic insight. One question that remains is what is meant by 'examination' when one says that Socrates, or any other questioner, examines a view by subjecting it to discussion and critical questioning. Aristotle had something to say about this matter, as shown in section 3 below.

To summarize then, Plato did not really define dialectic in a precise way that would be useful for logic to adopt. But by presenting actual dialogs on philosophical issues, he offered the best examples or paradigms that have ever been given. These dialogs, as shown by Robinson (1962), exhibit four rules that are characteristic of how the dialogs work. The first rule is that dialectic is essentially a social process in which two parties collaboratively take turns questioning and answering. Second, the respondent must reply to a question by giving his real opinion. But that opinion is expected to change, or be modified during the dialog, and a subrule is that the respondent is expected to maintain consistency in his opinions. The third rule is that the respondent must answer each question, and cannot just express ignorance. The fourth rule, related to the first one, is that the respondent cannot just defer to authority, but must try to express his own opinion. Beyond these rules, Plato does not tell us what the purpose of a dialog is. But one hypothesis is that the purpose of a Platonic dialog is maieutic in nature. It was left to Aristotle to give a more general and more precise account of dialectic that could be useful for developing dialectic as an abstract structure or formal method.

---

3.   In the *Theaetetus* (148a-151e) Socrates compared his method of philosophical discussion to the art of midwifery. As he tells Theaetetus (149a), his mother Phaenarete was a midwife.

3   Aristotelian dialectic

Aristotle tells us in the first sentence of the *Topics* that the purpose of dialectic is "to discover a procedure whereby we shall be able to reason about any problem set before us, from generally accepted opinions, and in our turn, stand up to others without self-contradiction" (*Topics* 1a). This statement tells us quite a bit about dialectic in the Aristotelian view. Dialectic uses reasoning to solve a problem. And it enables a participant to stand up to the other participant while avoiding self-contradiction. It seems then that in dialectic, self-contradiction is regarded as a bad thing or fault to be avoided if possible. This feature is familiar from Platonic dialectic. When Socrates shows that a given view[4] leads to a contradiction, this finding seems to be taken as sufficient to show that the given view must be wrong. But in one important respect, the Aristotelian conception of dialectic stands out as notably different from other Greek views of it, because of its stated emphasis on one special feature. The Aristotelian conception of dialectic is especially noted for its defining of dialectic with reference to its use of generally accepted opinions as premises. Aristotle, in *On Sophistical Refutations* (165b4), defined dialectical arguments as "those which, starting from generally accepted opinions (*endoxa*), reason to establish a contradiction". *Endoxa* are opinions generally accepted as having some standing by the mature and informed public and the experts. *Endoxa* are not just popular opinions of any sort. Nor are they necessarily the same as what is now commonly called public opinion, a term that can have many meanings. In the *Metaphysics* (1214b29–1215a15), Aristotle wrote that we do not have to "consider the views of the multitude" for they "talk without consideration about almost everything, and most about happiness". In many instances, a popular opinion, of the kind nowadays so often reported in a poll, may simply reflect the immaturity or lack of experience of the respondents. An *endoxon* is an opinion not only widely held, but that has some standing, and represents a serious view, or at least one that should be taken seriously. The word *endoxon* was translated as "reputable opinion" by Barnes (1980, p. 500). McAdon (2001, p. 124) agreed that although an *endoxon* can be characterized as a "generally held belief", it also needs to be seen in the Aristotelian sense as a belief or opinion that is "held in high esteem" or "of high repute". Renon (1998, p.95) has indicated that the adjective *endoxos* is somewhat ambiguous or equivocal. As applied to persons or cities, it can be translated as 'renowned', 'illustrious' or 'famous'. But as applied to views, opinions or beliefs, it

---

4.   The term 'view' is used in a technical sense here comparable to the notion of viewpoint (standpoint) defined by van Eemeren and Grootendorst (1994). According this technical sense of the term, a viewpoint is made up of a statement (proposition) and an attitude (pro, contra or neutral) towards the statement as expressed or advocated by a participant in a dialog.

has a more technical use in Aristotle's *Topics* and *On Sophistical Refutations*, meaning that a statement, proposition or premise is plausible to a greater or lesser degree. To say a proposition is plausible is to confer a kind of approval on it, saying that it is acceptable because it seems to be true, or that it is likely to be true. Aristotle defined *endoxa* in the *Topics* (100b22–24), as those opinions that are "reputable" and "accepted by everyone or by the majority or by the wise." This account is compatible with the view that *endoxa* can change, or can even be inconsistent with each other. Barnes (1980, p. 503) and Bolton (1990. p. 197) have pointed out that the opinion of most men can conflict with the opinion held by " the wise" or "the most reputable of the wise" (the experts), and that this means that one *endoxon* can be logically inconsistent with another. As Barnes (1980, p. 503) and Bolton (1990. p. 197) note, there is the possibility that the opinion of most men can conflict with the opinion held by " the wise" or "the most reputable of the wise" (the experts). As Bolton (1990, p. 197) points out, Aristotle was well aware that *endoxa* can conflict. Despite this potential for conflict, however, dialectic also presupposes some commonality of agreement in opinions.

But defined in this way, the concept of a dialectical argument seems useless, and even alien to the modern reader. For one thing, if an *endoxon* can contain a contradiction, and if a contradiction is a sign or error, what worth is an *endoxon* as a premise to start with? For another thing, the modern reader tends to assume that the purpose of a useful argument is to prove that a claim is true based on good evidence. By this criterion, reasoning from a generally accepted opinion to a contradiction seems to be useless. How could that get us to the truth of a matter by proving a claim based on good evidence? But perhaps there is a way, based on the notion of indirect proof or *reduction ad absurdum* attributed to Zeno (see section 1, above). By exposing false, questionable or indefensible opinions, an arguer could perhaps get along part way towards the truth of a matter being discussed. Thus the notion of the *endoxon* links naturally to the notion of a philosophical dialog or discussion in which commonly accepted views are subject to critical questioning by using logic to infer consequences from them. The examples that Aristotle would have had in mind are of course the Platonic dialogs. Socrates takes as hypotheses or starting points the reputable opinions of those who profess to know the truth of some controversial matter. He then uses question and answer sequences combined with careful logical inferences to draw out the logical consequences of the initial hypothesis. In the pattern that is so familiar in many of the dialogs, the line of argumentation may even arrive at a contradiction. In other ways, Aristotelian dialectic also resembles the Platonic concept of dialectic embodied in the Socratic dialogs. It is a process of questioning and answering. It involves (centrally) two people. It is based on the opinion of the respondent. And it involves a kind of rational persuasion whereby the proponent convinces the respondent to come to accept

some proposition he did not accept beforehand. As Evans (1977, p. 75) described Aristotelian dialectic, "success is achieved when one has secured the agreement of a particular opponent", and "to secure this agreement one must produce a sense of conviction" in that opponent. In these respects, the dialectic of Aristotle is similar to that of Plato.

Perhaps the best idea of what Aristotle takes to be the working method of dialectic can be gotten by examining his book the *Topics*. In the *Topics* (101a5–101b4), Aristotle tells us that the method helps speakers see multiple sides of an issue, enabling them to tell the difference between truth and falsehood on a certain matter in a better way. In other words, what the method helps us to do is to be better able to weigh the strengths and weaknesses of an existing argument on a controversial issue in a better way than we would be able to do without the aid of the method. For example, the method can be used to refute an *endoxon* by showing that it leads by logical reasoning to a conclusion that is antithetical to other propositions previously accepted in a dialog. Another important use of dialectic Aristotle also mentions is that it can help us to investigate the basic principles (*archai*) of a science by investigating them in light of existing *endoxa* about them.

As an example of the former use of dialectic, one might cite the passage in the *Nicomachean Ethics* (1096a5–1096a9) where Aristotle refutes the opinion that happiness corresponds to the life of making money.

> The life of money-making is one undertaken under compulsion, and wealth is evidently not the good we are seeking; for it is merely useful and for the sake of something else. And so one might rather take the aforementioned objects to be ends; for they are not loved for themselves. But it is evident that not even these are ends – although many arguments have been thrown away in support of them. Let us then dismiss them.

Here Aristotle refuted the opinion that happiness corresponds to the life of making money by arguing that while money is loved for the sake of something else, happiness is loved for itself. The proposition that happiness is loved for itself is taken by Aristotle to be previously accepted in the dialog that has occurred already at this point. It is a kind of *endoxon*, one that can be supported dialectically by other arguments.[5]

Where Aristotle's account of dialectic is different from Plato's is that, while it also takes the Socratic dialog as a paradigm, it generalizes from this particular exemplar, using its main features to construct an abstract model into which many different kinds of dialogs can fit. The device that makes this generalization possible is the *endoxon*. What is important in Aristotelian dialectic is that the view being

---

5.   Sara Rubinelli provided helpful discussions that shaped my views on Aristotelian dialectic, and drew my attention to this example.

examined in the dialog is not just the personal opinion of a particular opponent. But according to Aristotle (*On Sophistical Refutations* 170b5–170b8), what is important in studying sophistical refutations is not what makes them convincing to chance individuals. What is important in the study of fallacies, or apparent refutations (*sophistici elenchi*) is how they seem plausible as arguments to the general public, or the typical arguer who would be influenced by such an argument.[6] Thus the view being examined in the framework of Aristotelian dialectic must be more than the idiosyncratic belief of the particular respondent in the dialog. It must also be a "reputable" view that appears plausible to a wider group as a position that is worthy of being taken seriously on some issue of concern, or that is controversial. Thus the dialog that examines this view is of interest, and even enlightening, not just to a single person. The dialog is interesting because it explores the viewpoint itself, finding the reasons that can be given to support it, and those that can be given to refute it, or reduce its plausibility. As Evans (1977, p. 84) noted, in the *Topics*, Aristotle frequently appeals to "what seems to be so", not in the sense of what seems to be so for some particular individual, but in the sense of what generally seems plausible in the sense of being an *endoxon* or reputable view. Thus an Aristotelian dialog, to be successful, needs to be based not just on the individual belief of the respondent. It needs to be based on the commitment of the respondent to a viewpoint that is an *endoxon*. Otherwise the dialog could conform to all the rules of dialog, but it would not show us anything enlightening, or lead to discovery. Such a dialog could be formally correct, in following all the rules for questioning and answering. But it would not be useful in the sense of fulfilling its purpose as a dialog. In other words, a distinction needs to be made between the goals of each participant in a dialog, and the goal of the dialog as a whole collaborative structure of activity to which the participants are contributing.

But this account of dialectic raises another question that is continually worrisome in the modern endoxic view of argument and reasoning. If dialectic starts from premises that are only opinions, and sometimes end in contradictions, showing that the original opinion is not true, what is shown is that this form of reasoning does not always prove that the conclusion arrived at is true. It appears to be a

---

6.   *Endoxa* are not views that are universally accepted. But neither are they views that are just accepted by one individual, the respondent in a dialog. Evans postulated of Aristotelian dialectic (1977, p. 82), "A view which is universally accepted could not form a subject for dialectical debate." An *endoxon* is somewhere in the middle. It will be shown in chapter 3 how the Aristotelian notion of the enthymeme is connected to Aristotle's dialectic. Enthymemes are associated with generalizations that are not universally true, but only true "for the most part" or with qualifications, and are subject to exceptions. The connection with the study of sophistical refutations, or fallacies as they are now called, will also prove to be very interesting in later chapters in getting a grasp of Aristotelian dialectic as a whole.

fallible form of argument. It could support a claim, but then later on in the discussion, it could turn out to refute that claim, or show that it is dubious.[7] Shouldn't logical reasoning be firm in proving or disproving a claim? Is it trustworthy if what is proved now needs to be retracted later? There is an answer to this question that can be found in Aristotle's views about deliberation in ethical and political cases where there is inherent uncertainty and variability of circumstances.

In his writings on ethics, Aristotle took the view that in reasoning about things that are variable, "we must be content with premises to indicate the truth roughly and in outline", and with reasoning based on generalizations that are "only for the most part true" (*Nicomachean Ethics* 1094, 19–23). Dialectical reasoning can be useful, in his view, even if it is based on fallible premises that might turn out not to be true, as our knowledge changes. His view of scientific reasoning, or what he called "demonstration", in contrast, was that it must be based on axioms that are not variable, premises that are known to be true.[8] But trying to use demonstration in cases where the evidence is highly variable could be a kind of error. It is like assuming that because someone is a more skilled scientist or mathematician, he will also be a better administrator, political leader, or judge. This view of Aristotle's is necessary to fully grasp, in order to gain any appreciation of what Aristotelian dialectic is all about. But it will not be until chapter 3, where dialectic and rhetoric are compared, that its import can be fully realized.

Both Plato and Aristotle were careful to distinguish between dialectic and eristic or agonistic dialog, in which the sole purpose is to get the best of the other party by any means. It is only by means of this fundamental contrast that we can begin to get some grasp of what dialectic was meant to be as an art in the ancient world. According to Plato (*Sophist* 231e), eristic and agonistic are the arts of the sophist. In many passages in the Platonic dialogs, the sophist is portrayed negatively as an arguer who has no regard for the truth of a matter, and who is motivated by fame and money. According to Aristotle (*On Sophistical Refutations* 171b23–171b29), sophistic only appears to do what dialectic really does. Dialectic really does test the claims of those who profess to know, by using rational argumentation to critically probe the claims tested. But sophistic is a quarrelsome or eristic type of dialog where the goal is to simply attack and defeat the opponent by any means, fair or foul. In the eristic type of dialog therefore, no conclusion can be drawn about which side really had the strongest justification to support its claim. Plato used 'eristic' as a term of abuse, and associated it with antilogic (Nehamas, 1990, p. 8). Both Plato and Aristotle were clearly worried about misapplications

---

7.   The Aristotelian roots of the study of such defeasible inferences, and the defeasible generalizations they are based on, will be explained in chapter 3.

8.   Hamblin (1970, p. 16).

and abuses of dialectic, in the forms of antilogic and sophistic. The potential power of dialectic as a method of getting toward the truth of a matter was counterbalanced by the potential of misusing the same skills to create confusion and mischief. Dialectic was, especially for Aristotle, always tied closely to the subject of fallacies or sophistical refutations. In dialectic, real refutations need to be used to challenge an argument and throw it into question. But sophistical refutations can look like real refutations, and thereby be used to fool the uncritical onlooker or participant. Here once again, the point needs to be made that what was important for the usefulness of Aristotelian dialectic was a notion of persuasion based on what seems to be true, not just to the individual person with whom one is engaged in a dialog, but to the endoxic arguers who generally advocate the reputable view being discussed. The notion of what seems generally plausible is important.[9]

Beyond this brief sketch, there are many points of disagreement about exactly what Aristotle meant by 'dialectic'. Devereux (1990) commented that despite Aristotle's having written an entire book on dialectic – the *Topics* – the real function or role of dialectic in his philosophy is unclear: "The array of views on this question held by recent commentators is nothing short of bewildering."(p. 264). Despite this lack of agreement, and despite the unfamiliarity of the perplexing and even alien notion of dialectic to the modern mind, a kind of general impression of what dialectic was broadly taken to be in the ancient world does emerge. It is an art of rational discussion in which a questioner and a respondent reason with each other by question and answer. It contains arguments, and chains of connected steps of argumentation running through the sequence of dialog. It is adversarial but also partly collaborative, and hence different from mere quarreling, as in eristic or sophistical argumentation. It explores or probes into a controversial topic, very often of ethical interest, as illustrated by the Socratic practice portrayed in the Platonic dialogs. It is based on an initial opposition or disagreement between the two sides. Its premises are based on the respondent's answers, and also on generally accepted opinions. It is critical in nature, and proceeds by finding defects, weaknesses, contradictions and fallacies in the arguments, as they are developed through the dialog. It sometimes ends in perplexity, without resolving the initial conflict of opinions decisively, one way or the other. But even so, it seems to have a kind of therapeutic benefit, or gain in learning, through clarification of ambiguities and the testing of plausible or implausible assumptions through discussion. Its aim seems to be to move toward the truth of the matter discussed, and the Socratic

---

9.   As Evans (1977, p. 85) argued, plausibility can take an absolute or a relative form the way Aristotle discusses these notions in the *Topics*, in connection with the various topics, or forms of argumentation. Relative plausibility represents what is plausible to some given person. Absolute plausibility represents what is generally plausible, not just to one person, but to everyone or the majority or the wise. Absolute plausibility is closely related to the notion of the *endoxon*.

dialogs illustrate how this aim can be fulfilled. But there is also a dark side to dialectic. It can degenerate into antilogic and eristic, and be abused by clever sophists who use it to deceive.

Despite this broad idea of Greek dialectic that does emerge as a coherent notion, it remains hard for the modern mind to appreciate the significance or importance of dialectic the way the Greeks evidently did. The modern view is that merely talking about something, even though it could be therapeutic psychologically, will not in the end decisively prove anything, or establish what is true or false. Aristotle himself contrasted dialectic with what he called 'demonstration', or scientific reasoning from premises known to be true. Perhaps dialectic has been ignored or devalued in the modern view because is based only on opinion, and is therefore classified in modern terms as "subjective" rather than "objective". It seems to be a distinctively modern view that dialectical opinion-based reasoning, because it is admittedly subjective in just the way Aristotle says, has no proper place in logic, which is an objective and precise scientific discipline. This modern viewpoint could be the reason why it seems so hard for many of us now to appreciate or even grasp the Greek notion of dialectic.

Although Plato characterized dialectic in a narrower way by portraying it as a philosophical discussion of the kind exemplified in the Platonic dialogs, Aristotle made dialectic into a much more general method. In *On Sophistical Refutations* (179a20–179a36) he argued that scientific demonstrations depend on knowledge of a particular domain, like geometry, whereas dialectical arguments are common to every art and faculty. As Smith (1993, p. 338) pointed out, a feature of dialectic that is of great importance for Aristotle is that it is completely general in application. Thus, for Aristotle, dialectic is not just restricted to philosophical discussions, and is much more general. It could be applied to all kinds of dialogs, like everyday conversational exchanges, or any debates or arguments where there are differences of opinion. Indeed, there are some passages in Aristotle that can be taken as an attempt to classify the different types of dialog characteristic of dialectic.

4   Aristotle's classification of types of dialog

What is most interesting about Aristotle's theory of dialectic is that he classified different types of reasoning and different types of argument, and this system of classification could be seen as the first attempt to distinguish the distinctively different types of dialog. In the *Topics* (100a27 – 101a4), Aristotle classified three types of reasoning: demonstration, dialectical reasoning and eristic (contentious) reasoning. But he also classified several different types of argument. As indicated in chapter 1, this distinction can be taken as the first known systematic attempt by a

philosopher to make a systematic classification of the types of dialog important for dialectic. In *On Sophistical Refutations* (165a40 - 165b12) Aristotle classified four types of arguments: didactic arguments, dialectical arguments, examination arguments, and contentious (eristic) arguments. Didactic or pedagogical arguments are those used in teaching. Dialectical arguments, as indicated above, are defined as those which, "starting from generally accepted opinions (*endoxa*), reason to establish a contradiction" (165b4-5). Examination arguments (*peirastikoi logoi*) are "based on opinions held by the answerer and necessarily known to one who claims knowledge of the subject involved". Examination (peirastic) arguments are used by a proponent in a dialog to test out the views of the respondent by subjecting them to critical scrutiny. Contentious arguments (*eristikoi*) are defined as ones that "reason or seem to reason from opinions which appear to be, but are not really, generally accepted" (165b8-10). Contentious arguments are, of course, identified with sophistry and quarrelling. Aristotle's classification of these four types of arguments, or uses of arguments, could be held to be the basis for calling him the founder of dialog theory as a systematic subject, even though he did not himself invent the subject. And yet his remarks have proved to be nothing short of perplexing.

What peirastic arguments are, in particular, seems to be a mysterious question. Hamblin (1970) pointed to ambiguities suggesting that Aristotle may have sometimes held that the view that dialectical and examination arguments are in the same category. According to this view, examination arguments could be subspecies of dialectical arguments, and so there would only be three basic categories. Hamblin (1970, p. 59) cited the passage in *Topics* (159a25), as evidence for this view. In this passage, Aristotle contrasted examination arguments with contentious arguments, but there does not seem to be a clear distinction made between dialectical and examination arguments. Guthrie (1981, p. 155) also saw peirastic as being part of dialectic, or tied in with it, but appeared to see it as somewhat distinct as well. These issues have in the past largely been treated as trivial and obscure, except by some commentators like Guthrie who have seen something of interest in them. But they in fact become quite important in relation to recent developments in computing, as will be shown in chapter 5. The issue for dialog theory is how peirastic dialog should be analyzed and classified. Is it a species of information-seeking dialog? It does seem to be at least partly. For when you are examining somebody, presumably you are trying to get information out of them. But there is also a testing element involved. Examining seems to be kind of probing in which the information is tested out. Perhaps then examination dialog is best seen as a blending of information-seeking dialog with some other kinds of dialog, like persuasion dialog or inquiry.

It could be that examination dialog is of several different kinds. In his commentary of Aristotle's classification of the types of dialog, Guthrie (1981, p. 155)

saw a distinction made between peirastic argument and a related type of argument he called exetastic. According to Guthrie, peirastic argument involves a kind of dialog that could be called "testing or probing", while exetastic argument involves a kind of dialog that could be called "examining critically" (Guthrie, 1981, p. 155). But what does Guthrie take to be the difference between these two types of dialog? It's hard to make a guess. But the way Guthrie describes peirastic dialog makes it sound like information-seeking dialog, even though the purpose of getting the information from the respondent might be to test his knowledge. The exetastic type of dialog is more critical. Part of its purpose is to criticize a respondent's arguments. Thus exetastic argument seems closely related to dialectical argument. Of course, you might say that the peirastic type of argument is testing or probing, and may also involve some elements of criticism and dialectical argument. But the peirastic argument, as contrasted to the exetastic, appears to be even more deeply dialectical in nature.

It is hard to know how Aristotle really meant to classify examination dialog, and how much weight can be put on the distinction between peirastic and exetastic dialog. But the hypothesis that exetastic argument can be typified as having a critical aspect was put forward in the ancient manual on rhetorical argumentation called the *Rhetorica Ad Alexandrum* (1427b12-1428a17). *Exetasis* is described in this manual as a critical type of dialog in which the proponent attacks the respondent for "not practicing what he preaches". Exetastic argument is defined as an argument where the respondent's words or acts are criticized on the grounds they are contradictory to one another or even to the repondent's mode of life (1427b13-1427b14). By this account, the exetastic type of argument corresponds quite well to the descriptions of the circumstantial *ad hominem* argument given in modern logic textbooks. The proponent uses this kind of argument when he gets evidence that the respondent is committed to some kind of personal inconsistency, and then uses that against him to attack his argumentation.

Examination dialog turns out to be quite interesting for a number of reasons. The peirastic type of examination dialog could represent the familiar kind of dialog in which teaching takes place. It can be defined as a special form of information-seeking dialog in which the teacher already has the information, but her purpose is to find out whether the student also has the information. Of course the usual means of finding this out is the oral or written examination (test). Of course, in examination (peirastic) dialog the questioner already has the information and so it is not the usual kind of information-seeking dialog in which the questioner lacks the information and wants to get it from the respondent. Another familiar type of examination dialog is the kind of examination of a witness practiced in court. Here, the probing and testing tends to be more critical. This type of examination could perhaps be seen as exetastic. It is hard, at this point, to make any final

pronouncements on what the exact difference between the peirastic and exetastic types of dialog should be taken to be. In the sequel, we will adopt the policy of using the generic term 'examination' to refer to this type of dialog, and use the expression 'peirastic dialog' as being roughly equivalent to examination dialog. One can see from considering Aristotle's brief comments on these types of dialog that their importance was known in the ancient world. But exactly how peirastic dialog should be defined remains something of a mystery. It seems to be basically an information-seeking dialog, but perhaps also extended by a persuasion dialog component that gives it a critical edge.

While Plato gave us the examples of Socratic philosophical dialog that are the originating paradigm of dialectic, Aristotle defined a broader concept of dialectic that made it potentially applicable to all argumentation not only in academic disciplines but in education, law, politics, other professions, and even in everyday conversational exchanges. This widening of generality was also accompanied by an attempt to define dialectic in a more precise way that potentially has the structure of a rule-governed activity, or even a potentially formalistic framework or set of organized activities. Both the Platonic and Aristotelian models of dialectic shaped further developments of the subject in the Middle Ages.

5   Medieval dialectic

Logic as taught in the Middle Ages was an extension of the Greek model. Syllogistic was very prominent, but Greek dialectic also survived in at least one form. The form of teaching by question and answer has even continued through the Middle Ages into the modern world, in the tradition of the oral defense of the "thesis" in universities. The Aristotelian tradition of having the study of fallacies as part of the logic curriculum also continued. But in addition to these standard Greek features, there were some additions and modifications to Greek dialectic. There were even attempts made to systematize dialectic, by setting it up as a regulated game that formalized the various elements into a well-defined structure. The most prominent attempt of this sort is the so-called game of Obligation, a kind of question-answer game with definite rules. The way the game was set up, it could be played like any sort of game, like Monopoly. This aspect of the game of Obligation made it potentially attractive as a pedagogical tool, or perhaps even as a sort of measurable test of dialectical skill.

Many treatises on Obligation were written in the Middle Ages. An early one is William of Sherwood's *Treatise on Obligations*, which according to Stump (1989, p. 7), was probably written by Water Burley. Obligation evolved into different forms in treatises written by later writers like the fourteenth century logician Roger

Swyneshed. There have been many guesses about the purpose of these treatises. According to Stump (1989, p. 179) they are supposed to represent disputation in a stylized and structured form of a kind derived by the scholastics from Aristotle's *Topics*. According to Sherwood's account of Obligation, the one side in the disputation begins by advancing a proposition that he is then obliged to defend. The second party then proposes a number of propositions, one at a time, which the first party must accept, reject, or classify as doubtful (Stump, 1989, p. 180). The second party uses a sequence of such questions to try get the first party to commit himself to a contradiction. The first party tries to avoid contradiction. The games are interesting because they show that there is a formalistic aspect of Greek dialectic. The rules of Platonic dialog cited above suggest that one could devise a game with such a set of rules, and formally structure it as a zero-sum game with a winner and a loser. This suggestion is interesting because if dialectic can be formalized, it could be developed into some kind of formal logical system that could have many applications. It could be used to study fallacies, for example. It took a long time for such a program to actually be implemented as a serious research proposal. Finally in his book *Fallacies* (1970), Charles Hamblin worked out a systematic program of this sort, based on formal models of dialog meant to be applied to the study of fallacies.

The formal features of the game of Obligation have been well described by Hamblin (1970, p. 260–264). The main features of the game can be explained by using a simple and common form of example. Obligation is a kind of game, meaning that is has clear and precise procedural rules. In particular, it has clear win-loss rules. There are two participants. They were usually called the Opponent and the Respondent (Hamblin, 1970, p. 260). But for the purpose of our illustration, let's call them the Questioner and the Respondent. The Questioner has the right to ask yes-no questions to the Respondent. The Respondent is obliged to give an affirmative or negative answer. At the beginning of the game, the Respondent has to accept some designated statement as his thesis. It tends to be a statement that is known to be false. For example, the Respondent may be assigned the thesis, 'Socrates is black.' The Respondent's goal is to avoid self-contradiction. The Questioner's goal is to get the Respondent to contradict himself. When that occurs, the game is over. But if the Respondent can survive the dialog for a designated number of moves without contradicting himself, then he wins the game.

Suppose for example, that the Respondent has for his thesis the statement, 'Socrates is black.' Suppose the Questioner then asks, "Is Socrates the same color as Plato?" The problem is that, as the Respondent gets further and further away from the truth, it gets more and more difficult for him to avoid falling into unanticipated inconsistency. But since he is already committed to the statement 'Socrates is black', to maintain immediate consistency, he is now faced with a dilemma. If he

answers 'yes', he is now committed to the false statement that Plato is black. If he answers 'no', he is committed to the false statement that Socrates and Plato are not the same color. Such outcomes are potentially confusing, and could potentially lead to inconsistency down the line, as the sequence of questioning proceeds. The Respondent, to survive the questioning and win the game, must try to remember his prior replies and stick to his original thesis, without unwittingly committing himself to an inconsistency, given that his commitments may be diverging more and more from familiar reality. This formal approach represents the attempt to deal with the basic problem of dialectic posed by Plato. It is the problem of how to manage commitment in a dialog by balancing two of the Platonic rules. One is the rule that the respondent must give only his real opinion in answer to a question. The other is the rule that the respondent should maintain consistency with his previous answers (and, of course, with his initial thesis to be defended).

The advantage of the Obligation game, or of similar games of dialectic, is that, by having such clear rules, it is amenable to formalization. Hence, as Hamblin showed, the medieval games of dialectic were the precursors of modern attempts to develop formal dialectical systems, tying ancient dialectic to the modern attempts to study argumentation and fallacies. The disadvantage of Obligation is that such a game can appear to be trivial and mechanistic. The win-loss rules even make it appear to be agonistic, again making it appear more like what Plato and Aristotle would have called eristic dialog. This view of dialog is not in the collaborative spirit of discovery that made the ancient Greek dialectic appear potentially as such an enlightening process of self-discovery and intellectual refinement of views that really mattered. Unfortunately, medieval dialectic tended to be largely unappreciated after the Renaissance, and especially after the Enlightenment period. Typically it has been ignored, or if any attention is paid to it, dialectic has been portrayed as an instance of sterile and trivial logic-chopping typical of the Dark Ages. Especially with the rise of science in the Enlightenment period, the shift to a new conception of rationality brought with it a distrust of medieval reliance on Greek authority. Now empirical and mathematical science became the model of rational thinking. Dialectic was abandoned as a part of logic. Aristotle's deductive logic, the theory of the syllogism, along with the Stoic propositional logic, in its deductive form, were taken to be the models of all that really mattered in logic. The study of fallacies was pushed to the fringes of the logic curriculum and classified as "informal". Logic was on a long road towards its flowering around the beginning of the twentieth century as a mathematically formalized discipline centering around deductive propositional logic and the quantification theory. Dialectic was rarely mentioned, except as a Marxist-Hegelian metaphysical notion. The ancient notions that it could be formalized, and used a serious tool to study fallacies, had long faded into neglect and obscurity.

6   Dialectic in modern philosophy

What characterized the Enlightenment was not only the development of modern science, but with that the equation of science with objective knowledge. Anything outside the methodology of modern science came to be seen as subjective. Scientific reasoning came to be equated with rationality itself. Time and time again, it was impressed on everyone as the new conventional wisdom, when science clashed with popular opinions, science turned out to be in the right. Enlightenment philosophers, like Pascal and Descartes, espoused geometric reasoning as the model. One had to start with axioms and then use deductive logic to prove conclusions. In his *De L'Esprit Geometrique* (1659) Pascal proposed that all reasoning should be based on self-evident axioms and deductive inferences that yield unshakable proofs. This idea was not terribly new, as geometry was well known to the Greeks, and Aristotle had long ago proposed the notion of demonstration as the model of scientific reasoning. What was new was the rejection of the worth of opinion-based reasoning, of the kind characteristic of ancient dialectical argumentation, as having any serious import as a model of rational thinking. With the advent of the Enlightenment view of rationality, dialectic was simply excluded as representing a serious kind of reasoning that could lead to the truth, or that could establish a conclusion as known to be true.

With the Enlightenment way of thinking came the assumption that what was important was knowledge, especially as exemplified by scientific knowledge. Opinion could now be relegated to the realm of the subjective, as opposed to knowledge, which was taken to be objective. The Greek notion of dialectic as a kind of endoxic opinion-based discussion proceeding by question and answer was now relegated to the realm of "idle chatter". Yes, you could discuss your opinions, but such discussions notoriously go on and on, without ever getting to the truth of the matter or "proving" anything. It became a problem to know what to do with dialectic. Given its evident importance in the ancient world, modern systematic thinkers had to find some place for dialectic. And they did. Leibniz saw dialectic as an art of controversies (Dascal, Racionero and Cardoso, 2006). Kant and Hegel used the old term 'dialectic' in different ways.

Kant divided his transcendental logic into two sections – transcendental analytic and transcendental dialectic. The transcendental analytic was taken in Kant's *Critique of Pure Reason* (B85) to represent the formal aspect of knowledge. It represents the form of what is known. The transcendental dialectic was taken by Kant to represent what is not known, but only thought to be known. For Kant, dialectic is a critique of illusion. Dialectic, according to Kant, is made up of antinomies, or fundamental contradictory theses, with plausible arguments on both sides. For

example, the first antinomy (*Critique of Pure Reason* A426) is made up of the following pair of theses.

> *Thesis*:   The world has a beginning in time, and is also limited as regards
>   space.
>
> *Antithesis*: The world has no beginning, and no limits in space; it is infinite
>   as regards both time and space.

In Kant's view, a "proof" or chain of plausible reasoning can be given that supports the thesis, but an equally plausible chain of reasoning can be given that supports the antithesis. What this shows, in the Kantian view, is that it cannot be known which proposition is true, the thesis or the antithesis. Nevertheless, we are tempted to argue for one side or the other, and the whole province of such tempting arguments is that of dialectic. Thus it seems that Kant viewed dialectic in a negative light. Dialectic is the province of the illusion of knowledge. But since dialectic goes beyond knowledge, it can never establish propositions as true or false. Kant, the advocate of the so-called critical philosophy, saw it as a great error to think that we can ever get beyond the form of knowledge, and start to treat of the content of knowledge. Trying to promote or maintain the illusion that we can talk meaningfully about the content of knowledge is what constitutes dialectic, according to Kant.

In keeping with his negative view of dialectic, Kant saw ancient dialectic as a sophistical art – a logic of illusion (*Critique of Pure Reason* B86):

> However various were the significations in which the ancients used the word 'dialectic' as the title for a science or art, we can safely conclude from their actual employment of it that with them it was it was never anything else than the *logic of illusion*. It was a sophistical art of giving to ignorance, and indeed to intentional sophistries, the appearance of truth, by the device of imitating the methodological thoroughness which logic prescribes, and of using its 'topic' to conceal the emptiness of its pretensions.

Such dialectic, according to Kant, is "mere talk", and is "quite unbecoming to the dignity of philosophy". Such a sweeping rejection, even vilification, of the ancient notion of dialectic, including the views of Plato and Aristotle, appears to be a strong dismissal of what was taken by the ancients to be a vitally important part of philosophy. And it is. Kant was an Enlightenment thinker. His program was to limit reason to what can be known, and to be "critical" about whatever pretends to be knowledge but goes beyond experience and the transcendental analytic.

Hegel, like Kant, had a negative view of Greek dialectic. In *The Science of Logic*, he called dialectic, as generally regarded in the past, an "isolated part of logic" which seeks to refute through "limited assertions" and "sometime has nullity for its result"(Friedrich, 1953, p. 193). According to Hegel, Kant "set dialectic higher"

by freeing it from its "semblance of arbitrariness" (Friedrich, 1953, p. 193). But even Kant's conception of dialectic, as expressed in his antinomies, was seen by Hegel as too abstract and negative (p. 194). To go beyond what he saw as such a negative conception of dialectic, Hegel developed a new positive view that dialectic should be seen as a "comprehension of the unity of opposites" (p. 194). Hegel went on to build a speculative philosophy around this new conception of dialectic. Hegel's dialectic appears to share at least some of the features of Greek dialectic. It is about opposition, and about the resolution of conflicts. But it also appears to be quite different from Greek dialectic in other respects. In its Hegelian version, dialectic is not tied to a verbal exchange between parties who take turns asking questions, putting forward arguments to each other, and trying to critically question or refute each other's arguments. Instead, it is no longer purely about sequences of speech acts, but about concrete reality, or what Kant called the "thing in itself". From conventional descriptions of it, how Hegelian dialectic works is that a thesis interacts organically with its antithesis, thereby giving rise to new knowledge though a process of synthesis. This dialectical process has often been called the Hegelian triad. But is the famous triadic thesis-antithesis-synthesis really an accurate account of Hegel's analysis of philosophical method?

According to Dove (1970), the much-cited triadic interpretation misrepresents Hegel's phenomenological method, and Hegel's *Phenomenology of Spirit* is better interpreted as being non-dialectical. Dove (1970, p. 622) thinks Hegel's method is better described as being descriptive than dialectical. According to Dove's interpretation, "Hegel's method is radically undialectical", because the process Hegel writes of arises out of the course of experience being described, not out of the thinking process of reasoning about that experience. It seems then that commentators are divided on the question of whether Hegel is a dialectician or not. It is widely assumed that he is, but Dove points out (p. 622) that "scores of commentators" have also denied that Hegel employed a consistently dialectical method. So there is a controversy about whether Hegel should rightly be interpreted as being a dialectician or not. Whatever the answer to this question is, certainly it has been widely assumed that Hegel's method is dialectical, as he claimed (or appeared to claim, judging from the passages cited above). This assumption led others to advocate methods that followed Hegel's and were taken to be dialectical.

Historically, Hegelian dialectic led to Marxist dialectical materialism. But the development of logic certainly did not follow Hegel's apparently dialectical method. Instead, there was a strong reaction on the part of analytical philosophy against what was perceived as the obscurity of Hegelianism, and formal mathematical logic was developed as the tool of choice for analytic philosophy. These historical developments led even further away from the original Greek conception of dialectic as a philosophical method. On the one hand, formal logic, developed from

Aristotle's syllogistic and the Stoic propositional logic, pushed the informal fallacies to the fringes of logic. Dialectic was no longer even considered as a serious part of logic. On the other hand, dialectic was seen as a Hegelian and Marxist concept that was important in speculative and political philosophy, but that had no place in logic or science.

By the approach of the second half of the twentieth century, dialectic appeared to have come to a bad end. Citing Aristotle's remark that dialectic is "useful for gymnastics" (*Topics* 101a28), Kapp (1942, p. 12) described dialectic as a "curious kind of mental gymnastics", presumably in which some kind of verbal sparring occurs. The process of verbal sparring is one in which two parties argue about a proposed problem and one tries to trap the other in self-contradiction. To modern conventional thinking, the closest approximation to this activity would be forensic debating. But to engage in it as part of logic, would appear to be bizarre. Such an apparently odd activity of logical training appeared to have no counterpart in the modern world. Dialectic was now officially dead. The idea that anyone could take something called "dialectic" seriously as an academic or scientific subject now had a huge burden of doubt to overcome. Now dialectic was, according to generally accepted opinion, associated with Hegelian-Marxist metaphysical and political doctrines. You would think, judging by the generally accepted opinions (*endoxa*) of those times, that any attempt to resuscitate it would be doomed.

7   The re-appearance of dialectic

Only quite recently have serious attempts been made to revive dialectic as a branch of logic. As shown in chapter 1, these attempts have come from not only logic, but also computing and communications. (Hamblin, 1970) constructed and used formal dialectical systems to study formal and informal fallacies. Other formal dialectical systems that could be used for this and other purposes, have been proposed by Hamblin (1971), Rescher, (1977), Hintikka (1979; 1992), Barth and Krabbe (1982), Mackenzie (1981; 1990), Carlson (1983) and Walton and Krabbe (1995). Empirical studies in linguistics have studied question-reply sequences in cases of everyday conversational exchanges (Schlegloff, 1988). Researchers in computer science (Silverman, 1992; Reed, 1998) are now using dialectical models to investigate problems in artificial intelligence. Recent developments in multi-agent technology require agents in a system to communicate with other agents in the system in order to carry out collaborative projects (Wooldridge and Jennings, 1995). Such multi-agent collaboration can only be carried out if the agents reason with other agents by presuming collaborative rules of conversation. This capability for collaborative conversational interaction requires what amounts to a dialectical

framework in which questions can be asked and answered. The need for such a framework in computational technology has even prompted some to write about the founding of the new field of computational dialectics. Much the same kind of dialectical direction has been taken in the field of pragmatics in linguistics, based on the work of Grice (1975), and in the field of communication studies, with the pragma-dialectical approach of the Amsterdam School (van Eemeren and Grootendorst, 1984; 1992). Developments within the field of logic have also resulted in a new dialectic (Walton, 1998), in which many different types of goal-directed dialog are classified and studied.

In a survey of recent work on dialog theory in philosophical and linguistic pragmatics, Dascal et al. (2005) showed that the linguistic study of speech acts has converged with a wider body of scholarship in linguistics in an interdisciplinary body of work studying dialog models of communication. Research on argumentation at the University of Lugano, for example, has revived the ancient and medieval tradition of the topics, building on recent developments in semantics and pragmatics within linguistics, to build a dialog-based approach to argumentation. Their model of argument intervention uses an argument generator for argument construction in strategic maneuvering during the argumentation stage of the critical discussion (Rigotti, 2007). One of the most important aspects of the model is that it takes communicative contexts into account by situating an argument in a context of communicative practices that include negotiation, mediation and deliberation (Rigotti and Rocci, 2006). In addition to an interpersonal dimension, there is also an institutionalized dimension that applies to argumentation through activity types as well as argumentation schemes. Through these studies it has been shown that an argument needs not only to be evaluated in relation to the normative model of dialog, but also in relation to how that normative model applies to factors arising from a specific activity type. For example, it might be scarcely possible to analyze and evaluate an argument that takes place in a courtroom without situating it in the context of the activity type of a legal trial. The reason is that a trial has specific rules, like rules of evidence and other procedural rules, that govern whether an argument is considered relevant evidence and if so, how the argument is to be evaluated according to standards and burdens of proof.

One of the most important concepts of argumentation to be considered in this light is that of common knowledge, a notion shown to be very important in the study of enthymemes in chapter 7. Realistic communication takes place against a background of knowledge, commitments and values shared by the participants. Argumentation to resolve a conflict of opinions between two parties cannot be successful unless the parties share some common knowledge in a dialog. There needs to be a set of statements that both parties can agree to as acceptable in advance of reaching the argumentation stage. The study of how arguments should be

evaluated in a dialog in a specific case relates to the common knowledge shared by the participants in that case. Typical examples of arguments are negotiated within a social context characterized by an institutional dimension that assigns rules to participants and influences who can say what (Rocci, 2006). Such dialogs have goals that are pursued alongside the goals of the normative model (like that of the critical discussion). According to the exponents of this approach, one that combines semiotics and dialog theory, both kinds of goals need to be taken into account when analyzing and evaluating arguments. Such contextual factors will turn out to be essential when we come to studying incomplete arguments by finding implicit premises or conclusions in them arising from common knowledge.

The natural and nowadays widely accepted way to classify dialectical systems is to define the goal of each type of dialog. However, the dialog systems in Hamblin (1970; 1971) are not presented in such a way that the central goal of the dialog is explicitly stated in a precise way. The motivating idea seems to be that the participants are arguing two sides of a disputation in which the one who has the strongest argument wins. Hamblin did say about some of the main formal systems he constructed that they were meant to be *information-oriented*, meaning that "it is assumed that the purpose of the dialog is the exchange of information among the participants." But Hamblin did not tell us precisely what he meant by "information". This omission can be seen as reasonable however, because the different types of dialog he pioneered were just at the beginning state. What appears to be the common pattern in Hamblin's systems of dialog is the evolution of an argument designed to persuade in small steps. Each party uses a series of single steps of inference to try to get the other party to become committed to some designated statement.

The dialog systems of Hintikka (1979; 1992), like Hamblin's systems, work by small steps of questioning and replying containing a thread of argument. The Hintikka system (1992) models Socratic elenchtic dialog of the type studied by Robinson (1962). Like the Hamblin systems, the Mackenzie formal systems appear to be based on a rational persuasion goal of dialog where each party uses a chain of argument steps to try to get the other party to become committed to the persuader's thesis. Most of the formal dialectical systems to date appear to be aimed at representing argumentation whose purpose is persuasion. However, it is possible to think of a different kind of dialog where the goal is to prove a designated statement, or disprove it, by collecting all the evidence in a methodical search. This type of dialog is called the inquiry in the typology of (Walton, 1998). Intuitionistic logic, in the well known format modeled by Kripke (1965), has been restructured by Barth and Krabbe (1982) to represent a dialog process of inquiry. A modeling of inquiry using a formal dialectical system has also been carried out by Felscher (1985). A survey of these various dialectical systems can be found in (Walton and Krabbe, 1995).

The classification of the six basic types of dialog given in (Walton and Krabbe, 1995) and (Walton, 1998) has been summarized in Table 1.2 in chapter 1. What centrally differentiates each different type of dialog is its distinct goal. An argument that may be used appropriately to fulfill one dialog goal may be inappropriate if used in a different type of dialog. What is revealed here is the multiplicity of different types of dialog. The typology in Table 1.2 is not meant to suggest that there are only six types of dialog. It is quite possible that there could be many other types, and also various combinations or mixtures of these types. But for purposes of evaluating arguments and other moves associated with informal fallacies, these six types recur, and have proved their practical importance. Given this typological system, a basis for classifying the various formal dialectical systems is possible. But little work has been done yet on seeing how the various systems fit together into some kind of unified academic field that could be defined as dialectic.

The multiple formal dialectical systems that have been developed are very significant, and show all kinds of potential, not only as interesting formalizations, but also as potentially applicable structures for studying informal fallacies and other matters of interest to informal logic. But because there is such a wide variety of technical systems, it may appear that dialectic lacks unity as a coherent field, or branch of logical and linguistic studies. The proliferation of different systems makes one wonder which system is the right one, or whether as technical constructions, all the systems are arbitrary. The answer to this question is to be sought in developing a definition of dialectic that brings all its technical features together, by setting out its most important or "essential" characteristics.

8   Eight characteristics of dialectic

In this paragraph, a brief explanatory sketch is presented, defining what dialectic is, in line with the ancient conceptions of it in Plato and Aristotle, but also in a way that could make it a potentially useful subject in light of modern applications. In the paragraphs below in this section, seven important characteristics of dialectic are described. Dialectic is a branch of the science or art that is now called logic. The primary purpose of dialectic is to provide a method for analyzing and evaluating argumentation used for some purpose in a conversational exchange between two parties. But dialectic could potentially have many other applications as well. For example, it could be used to analyze explanations, definitions, questions, and other parts of speech as well as arguments. In fact, dialectic could be useful for many purposes. But its primary aim should be seen as one of providing normative standards for the correctness and incorrectness of argumentation. The primary application of dialectic is to the study of fallacies. Dialectic is the method that

needs to be used to judge, in specific given cases, whether an argument can rightly be said to be reasonable or fallacious (or to fall somewhere between these two extremes). Emphatically it needs to be said that dialectic is an instrument of rational argumentation. Its primary goal is not how to actually persuade people to do something or accept something as true. Its goal is normative, or logical. Its goal is to provide standards for judging arguments as correct or incorrect, strong or weak, rationally justifiable or not. In this respect, dialectic is comparable to traditional formal logic. But how dialectic carries out this task is different from the way formal logic does. Dialectic, like rhetoric, views an argument as an exchange between two parties. And both subjects have to do with persuasion. But a key difference, as will be brought out in chapter 3, is the stronger emphasis of dialectic on the normative or logical aspects of argumentation.

The first distinctive characteristic of dialectic is that it involves two participants who are reasoning with each other. The contrast is with monolectical logic, which defines an argument as a set of propositions, one of which is designated as the conclusion. For dialectic, an argument is always an exchange between two parties, usually called the proponent of the argument and the respondent of the argument. As noted above, it is possible for one person to alternately take on the role of proponent and respondent in order to examine both sides of an issue dialectically. And of course it is also useful to mention that citing two participants is to begin with the simplest case. In principle, dialectic can involve three, four, or any finite number of participants. But the basic point is that the dialectical view of argument sees an argument as a claim made by a proponent and open to questioning or doubt by a respondent. Traditional formal logic, like the first-order logic of propositional calculus and quantifier logic, is monolectical. The argument is viewed as valid or invalid without essential reference to the arguer who makes the claim specified in the conclusion, or to any arguer who doubts that claim. No such references to claimants or doubters are required in the formal system.

A second important characteristic of dialectic is that it has to do with questioning and answering. According to the Stoic view, dialectic was characterized by the asking and answering of questions. As noted above, Diogenes Laertius reported that according to the Stoic view, "Without it [dialectic] it is impossible to ask and answer questions methodically"(Long and Sedley, 1987, p. 184). Plato was conspicuously aware of this aspect of dialectic, as shown above, and represented it at work in the Socratic dialogs. Socrates typically begins a dialog with a question. His respondent offers an answer, or attempts a reply. The whole sequence of questions and answers then typically leads to some problem or difficulty. Often the problem is a difficulty of a logical nature, like a contradiction between one of the respondent's current replies and one of his earlier replies in the sequence. Such a development is often called an elenchtic sequence of questioning, from the term

elenchos, meaning refutation, as explained above. Socrates used this kind of probing technique or elenchtic questioning to reveal logical puzzles, contradictions, or other shortcomings in a respondent's position, forcing the respondent to change his position. Such a change of opinion could represent an improvement, if the respondent was led to qualify or sharpen his view in light of the difficulty revealed by the questioning. The question-answer characteristic of dialectic makes it different from the monolectical logic we are more familiar with. Dialectic is not just about proving that a proposition is true or false. It is often about raising legitimate doubts and subjecting doubtful opinions to critical questioning and examination.

A third important characteristic of dialectic is that it is situational and pragmatic. It has to do with uses of an argument within a given conversational context. But central to dialectic is the aim of pursuing the truth of a matter being discussed. For example, in a Socratic dialog the purpose is to try to get towards the truth of a matter being discussed by using rational arguments to probe the issue and raise questions about it on both sides. Of course, dialectic may not be the only means of pursuing the truth. And many might say that dialectic does not need to actually arrive at the truth of the matter being discussed in order to be successful. It could be enough that it goes some way toward the truth by clarifying ambiguities and plausibilities, or by sharpening one's view through exposing logical problems or fallacies in it. But dialectic might not be exclusively concerned with Socratic discussions that try to get at the truth of a matter. There can be other kinds of conversational exchanges where the purpose might be different from that of a Socratic discussion. For example, two parties may deliberate with each other in order to make a choice between two possible courses of action. Or two parties may negotiate with each other in order to try to resolve a conflict of interests. There are many different kinds of goal-directed conversations. Each has its own purpose. An argument or move suitable to achieve the goal of one type of conversation might be quite unsuitable to move forward towards the goal of another type of conversation. Thus on an even wider view of its scope than was typical in the ancient world, which tended to emphasize philosophical discussion as the paradigm, dialectic could be seen as encompassing many different conversational contexts of argumentation. On a broader modern view, dialectic could be seen as also concerned with argumentation in conversational contexts like negotiation and deliberation, where using argument to try to prove a proposition is true (or false) is not the central conversational goal. According to this wider modern view, dialectic is relative to the type of conversational exchange that two parties in a given case are supposedly trying to take part in. On this view, the kinds of conversational exchanges involved could be multiple and varied.

A fourth important characteristic of dialectic is that, at the core of it, there is always an underlying opposition between the two parties. There is always some-

thing unsettled – a central disagreement, conflict, or difference of opinions between the two parties. The purpose is to resolve this disagreement by using rational argumentation. This opposition can be stronger or weaker in different cases. In some cases, one party advocates a certain opinion while the other part advocates the opposite opinion. In cases of weaker opposition, one party holds a certain opinion while the other party doubts that opinion, or has questions about it, even though he does not maintain the opposite opinion. But it is a mistake to think that dialectic is purely adversarial. Dialectic must have collaborative rules of polite conversation, of the kind cited by Grice (1975), even though dialectic also contains opposition. What dialectic requires is the right balance between advocacy, which is adversarial in nature, and procedural cooperation, which is collaborative in nature. The precise weighting of this balance needs to vary, depending on the type of conversation that the participants are engaged in.

A fifth important characteristic of dialectic is that to be significant, an argument does not have to be based on knowledge or belief. Instead, many arguments are based on acceptance, and that is enough. As Hamblin (1970) put it, the central notion of dialectic is commitment. As a participant takes part in a dialectical exchange, statements are inserted into or withdrawn from her commitment set. For example, if she asserts that a statement is true, that statement is inserted into her commitment set. If she retracts commitment to a statement, that statement is withdrawn from her commitment set. The commitment rules for a type of dialog determine which statements go in or out of a participant's commitment set each time he or she makes a particular type of move in that type of dialog (Walton and Krabbe, 1995). Commitment is not necessarily the same thing as belief, although it may often correspond with belief. Commitment represents what you have gone on record as accepting in a dialog.

A sixth important characteristic of dialectic is that it is based on premises that express generally accepted opinions. This social aspect of dialectic has often been overlooked by the modern theorists who have proposed the various formal dialectical systems. It is Aristotle who must get primary credit for recognizing the importance of it. As indicated above, *endoxa* or generally accepted opinions, are among the most important premises in dialectical arguments. Also, as noted above, it is possible on Aristotle's view that *endoxa* can contain or imply contradictions. By revealing such a contradiction, dialectic can be used to critically question and challenge accepted views. Thus an important public function of dialectic is revealed by the Aristotelian view. Apart from what is noted in the Aristotelian conception of dialectic, it is also possible for dialectic to get its premises from sources other than generally accepted opinions. Arguments in dialectic could also, in some instances, be based on premises that report information collected from a source.

A seventh important characteristic is that dialectic has a formal aspect as well as a descriptive, or applied aspect. As Hamblin (1970, p. 256) wrote, "The study of dialectical systems can be pursued descriptively or formally." Dialectic can be studied as a purely formal, mathematical discipline that investigates the formal properties of different kinds of structures appropriate for different kinds of conversational exchanges. It can also be studied from an empirical or descriptive viewpoint by examining texts of discourse of actual dialogs like parliamentary debates, trials, media editorials, talk show discussions, philosophical discussions, and so forth. The descriptive use of dialectic typically centers on a case study of some text of discourse containing argumentation. Special institutional factors may need to be taken into account. For example, in studying argumentation in a parliamentary debate, the setting and procedural rules may need to be taken into account. Or in studying argumentation used in a trial, the appropriate rules of evidence for the jurisdiction (FRE, 1997) may need to be taken into account.

The eighth characteristic of dialectic is that it has to do with the meanings of terms used in argumentation, and with the definitions that are put forward to represent these meanings. It especially has to do with words and phrases that have a strong emotive connotation that is either positive or negative, and are often linked to controversies, like ethical, political, and philosophical issues. This feature is clearly marked in the Platonic dialogs. In the *Meno*, for example, the dialog turns around various conflicting definitions of the term 'virtue' put forward for discussion by Socrates. This feature is also notable in the Aristotelian notion of dialectic. Although less frequently cited as a main characteristic of Aristotelian dialectic, it is taken to be very important by Evans (1977), who has a whole chapter on it (chapter 4). Of course, definition is clearly very important in the detailed discussions of the various forms of argument (topics) in the *Topics*, and in *On Sophistical Refutations* as well. But Evans (1977, p. 104) maintained that definition has an importance and interest in Aristotle's dialectic independently of these detailed discussions. In the modern literature, the dialectical importance of definitions can easily be appreciated by anyone who is familiar with the notion of the persuasive definition, developed especially by Stevenson (1944) as an important aspect of ethical, political and philosophical argumentation. A proposed definition, like any statement, can be put forward as a hypothesis in a dialectical discussion, and can be examined and questioned just as any other statement or opinion could be. Opposed definitions of controversial terms can be put forward and discussed as well. But since this feature is already so well exemplified in the Platonic dialogs, there is perhaps no need to try to illustrate it with further examples.

9   Hamblin's dialog rules

In Hamblin's chapter 'Formal Dialectic' (Hamblin, 1970, 253–282), he discussed a number of different rules, particularly rules regarding the asking of questions and the making of assertions, that could be appropriate for formal systems of dialog. Although the tone of Hamblin's discussion is very tentative, the potential rules he formulated are very interesting for a number of reasons. One reason is that they are quite general, and would appear to affect any kind of formal dialectical system that one might try to construct. Another reason is that the rules discussed clearly relate to various informal fallacies. Another reason is that many of the rules are quite perplexing, because, although they seem reasonable, it is hard to know whether they should be adopted or not. In several instances, it is clear that adopting a rule would be problematic. Another reason is that the discussion of these rules anticipates many of the fundamental problems about conversational policies in agent communication languages outlined in chapter 5.

   Before going on to examine these rules in detail, it should be noted that Hamblin is not very clear about the exact purpose of the formal dialectical system or systems which these various rules supposedly would be contained in. This vagueness is perfectly understandable, because the various types of dialog classified in chapter 1 had not yet been identified as such. Sometimes Hamblin writes as though the purpose of a formal dialectical system is to seek or exchange information. At other times it appears that what he has in mind is more like what would be classified in chapter 1 as a persuasion dialog. But this vagueness of Hamblin's discussion is not a bad thing, because it allows for a certain flexibility and generality in his discussion of how the rules might be used in different possible systems of formal dialog. It is clear that he is discussing the rules in a highly tentative way that is appropriate considering that they might be implemented in different ways in many different formal dialectical systems.

   One type of rule discussed (p. 268) concerns multiple questions of the form, 'Question A, B,...,C?', that are closed questions, meaning that they put forth a closed or finite set of alternatives, and the respondent (normally) would have to select one as the answer.[10] In Hamblin's set of syntactical rules for the answering of questions (p. 266), a "no commitment" option is allowed to the respondent. Following these rules, it would be considered an allowable reply for the respondent to make a non-committal reply to a closed, multiple question. He can reply that he is not committed to any of the statements A, B,..., C, or even that he is committed to the negation of each of them. So using this kind of rule could be a way of handling

---

10.   It might be noted here that we have changed Hamblin's terminology and notation slightly to conform to our style. For example, speaker and hearer become proponent and respondent.

loaded questions. But then Hamblin (p. 269) goes on to consider another way of solving the problem. He considers the following rule.

> *Question Rule 1*: Question 'A, B,..., C?' may occur only when A v B v... v C is already a commitment of both proponent and respondent.

This rule solves the problem by restricting the kind of question that the proponent can ask. He simply can't ask a loaded question of the kind that would force the respondent to agree to something he is not really committed to, or does not want to accept. But there is a problem with this solution. Hamblin comments that this rule would "seriously impede the asking of questions" (p. 269), and that a weaker rule should be considered. Basically the reason is that such a rule would allow the questioner insufficient latitude to probe into statements that are at issue in the dialog. For the respondent could simply always select as an answer the statement he is committed to, and the proponent could only pose alternatives where she (and the respondent) is already committed to one of them. Hamblin suggests that weaker variants of this rule might be considered (p. 269). For example, only the respondent might have to be committed to one of the alternatives.

The problem here can be seen as a kind of power stalemate between the two sides in a dialog. The questioner needs the power to ask questions that have presuppositions built into them that the respondent might not (yet) agree to. But on the other hand, the questioner should not have so much power that she can force the respondent to take on commitments that he really doesn't agree to, and would like to dispute, or remain doubtful about. Thus the respondent needs the power to reply 'No commitment' to statements or assumptions he is uncertain about, or doesn't think are justified. But if he has too much power to avoid commitment on anything he is queried about, he could prevent a dialog from making any progress at all. Question Rule 1 limits the power of the questioner so severely that she could hardly ask any interesting questions at all.

As well as rules governing the asking of questions, Hamblin also considered rules concerning the making of assertions. For example, the following rule is discussed (p. 269).

> *Assertion Rule 1*: 'Statement *A*' may not occur when *A* is a commitment of the respondent.

Assertion Rule 1 seems reasonable. For what's the point of making statement *A* if the hearer (respondent) is already committed to it? The statement performs no function of informing the respondent of something he does not know about already. As Hamblin put it (p. 269), this rule seems reasonable, "if we regard the sole function of statements to be the giving of information." Parenthetically, it might be noted that this rule foreshadows a similar rule concerning the speech act of in-

forming found in the agent communication languages in multi-agent systems discussed in chapter 5.

But then considering another question rule that would parallel this assertion rule, Hamblin (p. 269) goes on to formulate a rule that seems to be the opposite of Question Rule 1.

> *Question Rule 2*: Question *A*, *B*,..., *C* may not occur when any of *A*, *B*,..., *C* is a commitment of the proponent.

The rationale of this rule appears to be similar to that of Assertion Rule 1. This view of the function of questioning is called by Hamblin "the view of question as inquiry" (p. 269). Presumably what he referred to is the principle that there is no point in asking a question unless the questioner really wants to know about the statements she is questioning about (and so is not committed to any of them yet).

Hamblin considered two basic types of questions in his formal dialectical structures. The first type of question, which invites the respondent to select and answer from a closed list of alternative statements was illustrated in Question Rule 1 and Question Rule 2 above. The second type is called the why-question. The Hamblin type of why-question is basically a request for justification. When a proponent poses the question 'Why *A*?', she is requesting that the respondent provide a reason for her to accept (become committed to) *A*. Hamblin considers two rules that could be potential candidates for systems of formal dialectic (p. 271). The first rule concerns the asking of a why-question.

> *Why-question Rule 1*: 'Why *A*' may not be asked unless *A* is a commitment of the respondent and not of the proponent.

The second rule concerns the answering of a why-question.

> *Why-question Rule 2*: the answer to 'Why *A*', if it is not 'Statement not-*A*' or 'No commitment *A*', must be in terms of statements that are already commitments of both the proponent and the respondent.

The justification Hamblin suggested for Why-question Rule 1 is the comment, "Otherwise the 'why' is academic" (p. 271). This justification seems to make a lot of sense. For a proponent would only ask a why-question if the statement asked to be justified is in fact something the respondent holds, or is committed to. And the statement doesn't really need to be justified to the proponent if she already accepts it (is committed to it). The second rule also appears to make sense. If the respondent is going to justify a statement, surely the statements that he is going to use as premises must be commitments of the proponent. Otherwise they will not be useful to persuade the proponent to come to accept the statement that needs justification. At the same time, they should also be commitments of the respondent, if he

is going to use them as evidence to back up the statements he was asked to prove. Thus Why-Question Rule 2 also seems to make a lot of sense in the context of one party responding to a request to prove something to another party by giving good reasons that support it as acceptable. Not only that, Why-question Rule 2 successfully bans circular reasoning (Hamblin, 1970, p. 271). That would seem to be another strong reason for accepting it as a rule of formal dialectic. And yet, both these rules are problematic. Hamblin notes (p. 271) that Why-question Rule 2 "makes it impossible to develop an argument more than one step at a time." He therefore characterizes this rule as "unnecessarily strong" (p. 271).

The upshot of Hamblin's discussion so far is that once you try to formulate precise formal dialectical rules for the speech acts of questioning and asserting, different conflicting rules suggest themselves. But it becomes impossible to decide which is the right rule. Some rules sound right for one context, perhaps that of persuading, but seem inappropriate or problematic in a second context, like that of informing. On the other hand, the opposite rule might seem to serve well in the informing context, but then be problematic in the persuading context. What seems to be suggested is that asserting and questioning as speech acts do not admit of setting in any single dialog context of the use of argumentation. A rule that might be good in one type of dialog might distribute the balance of power between the proponent and the respondent quite problematically in another type of dialog. Even so Hamblin's approach looks promising, in that by setting up the rules on a commitment basis, he shows how there is great promise in formulating rules that are precise and reasonable in some context or other. It is just that the different types of dialog, like the distinction between persuasion dialog and information-seeking dialog, had not yet been introduced, refined and established.

10   Functions of questioning and asserting

According to Hamblin (1971, p. 137), all the formal systems of dialog constructed in his 1971 paper are "information-oriented". This remark suggests that they might be classified as information-seeking dialogs in the dialog typology system of chapter 1. It is a good hypothesis that they can be so classified, but Hamblin's remarks do not confirm it very fully. He did define an information-oriented dialog (1971, p. 137) as one where "it is assumed that the purpose of the dialog is the exchange of information among the participants." The formal systems of dialog constructed in (Hamblin, 1970) look quite similar in their general structure to those in the 1971 article, and once again Hamblin says very little about what the goal is supposed to be. But it would appear, although it is again only a hypothesis, that these could be classified in the typology of chapter 1 under the heading of persuasion

dialog. It is clear that in these dialogs, the respondent starts out by being uncommitted to some statement, and the efforts of the proponent are directed towards securing the respondent's commitment to this statement through the use of a connected sequence of arguments, one step at a time. In the 1971 paper (p. 148), Hamblin even describes the systems in the 1970 book by contrasting them with the information-oriented systems in the 1971 article.

> A system of interest which, however, is not strictly information-oriented would be one that permits a participant to develop an argument by securing assent to individual steps. Thus it might be held proper to ask questions whose answers could include immediate consequences of commitments, though not commitments themselves, and to demand "rationality" of participants in conceding such answers if they were not prepared to retract the commitments. I have elsewhere described an alternative argument-development system using questions of the form "Why?"(and here a footnote to the 1970 book is inserted).

This quotation is particularly interesting because it is as specific an account as Hamblin gives of the goal of the 1970 dialog system types, as contrasted with the goal of the information-oriented type of dialog represented in the 1971 systems. The 1970 type of dialog has two characteristics cited. One is the series of connected steps used by the proponent to secure the assent of the respondent. This characteristic suggests (at least partly) what is called persuasion dialog in chapter 1. The other characteristic of demand for "rationality" at least partly suggests the notion of burden of proof that is so important to persuasion dialog. It is not easy to grasp precisely what kind of dialog situation Hamblin is describing in this quotation. But it may be helpful to try to imagine a specific dialog situation. Suppose that the proponent in a dialog asks a question 'A?', and the respondent answers that yes, he is committed to A. But then suppose that B follows (logically and immediately, by one step of inference) from A. Now according to Hamblin's account of this type of dialog described above, the respondent's answer can now be taken to (indirectly) commit him to B. Now "rationality" can be demanded of the respondent. He can now be called upon by the proponent to make a choice. He can either concede B (even though he may not like B, and may not want to declare commitment to it), or he can retract his commitment to A. If this situation correctly corresponds to what Hamblin is describing, then it can be taken to express a notion of "rationality" connected with burden of proof that is characteristic of persuasion dialog. If so, a case can be made that Hamblin did recognize, at least partly, the distinction between persuasion dialog and information-seeking dialog.

   Another plausible hypothesis is that Hamblin's later book on imperatives could be taken to represent his explorations on yet a third type of dialog, which would be classified as deliberation dialog in the typology of chapter 1. Just below the quota-

tion cited above, Hamblin (1971, p. 148) wrote: "Of particular grammatical interest would be extension to some other kinds of locution, such as imperatives." This remark suggests that the later book, *Imperatives* (1987), could be seen as a move in a direction different from the other two types of dialog indicated above.

Just below his statement about dialogs being "information-oriented", Hamblin, (1971, p. 137) made another remark that amplifies what he meant by 'information' to some extent. This particular remark is worth quoting, because it is illuminating, with respect to certain issues about conversation policies in current multi-agent systems in computing that will be treated in chapter 4.

> It follows, for example, that there is no point in making any statement to someone who is already committed to it, or in asking a question when one is already committed to one of the answers. In practice statements sometime have other functions than to inform, such as to make an admission of something already admitted by others, or to exhibit the speaker's knowledge; and questions may serve as admission-elicitations or as knowledge-testing probes.

This quotation is interesting because it shows Hamblin recognizing the asking of a question and the making of a statement as localized kinds of actions that have a purpose or function as used in dialogs. In other words, Hamblin was on the verge of recognizing question-asking and asserting as communicative actions or speech acts, and of recognizing that these speech acts could have different functions or uses in different types of dialog. His remark foreshadows the later use of questioning and asserting as important kinds of speech acts in multi-agent communication languages. One can see that Hamblin was struggling with the very problems and issues that later came to be central to the development of communication methods in multi-agent systems. Here he enunciates some fundamental principles of communication.

The first principle could be called a general principle about the act of making a statement in a dialog, or as it would now be called, the speech act of asserting.

> *Statement Principle 1*: there is no point in making any statement to a respondent who is already committed to it.

But why should such a principle hold? What could be its rationale? This question seems really hard to answer, and probably that is why it has seemed so baffling, and continues to seem so even now, even though it seems plausible (reasonable), once expressed. The clue to demystifying the rationale of the principle is contained in Hamblin's insight that the act of making a statement can have different functions or uses in different types of dialog. First, a statement can be used to inform. This information function would presumably be found in an information-seeking type of dialog (or so we would now say). Second, a statement can be used to admit

something already admitted by others. This concession function would be typical of persuasion dialog, where the act of making a statement would often or typically be important as a way of making a commitment in response to the other party's persuasion argumentation. Third, the function of making a statement could be to exhibit the speaker's knowledge. This function is central in the special type of information-seeking dialog called the expert consultation type of dialog. In short, there can be three different ways of justifying and explaining Statement Principle 1.

Comparable to Statement Principle 1, a general principle applying to the asking of questions is articulated by Hamblin (1971, p. 137).

> *Questioning Principle 1*: There is no point in a proponent's asking a question to a respondent who is already committed to one of the answers.

This principle is quite general. It applies to all kinds of questions, including yes-no questions and why-questions. It also depends on what counts as an answer to a specific type of question, and in a specific type of dialog, it may be added. So how could such a general principle be justified? The question seems very puzzling, even unanswerable, until one recognizes that it may need to be justified differently in different types of dialog, recognizing that the asking of a question can have a different function in each distinct type of dialog. Two separate functions of the speech act of asking a question are cited by Hamblin (1971, p. 137). One is that of "admission-elicitation". Presumably this function is one that would occur in a persuasion dialog. As noted in chapter 1, it is very important for the question-asker in a persuasion dialog to try to secure the commitment of the respondent to some statement or other, as a premise in her chain of argumentation used for the purpose of persuasion. For this purpose, there is no point in the proponent's asking the question if the respondent has already indicated what his commitment is on the issue. Hence Questioning Principle 1 is justified in the context of persuasion dialog.

But could there be another type of dialog in which Questioning Principle 1 would have a different function, and therefore could be given a different type of justification? Intriguingly, Hamblin does briefly indicate such a function, called that of a "knowledge-testing probe". Presumably, this function would naturally occur in the context of the information-seeking type of dialog. The proponent wants to not only get information from the respondent, but also test out the worth or trustworthiness of that information. How does she do that? The answer is – by testing the information by following up with probing questions that verify it, or show you whether the respondent can back it up with evidence. The most familiar context of such a use, which could be called the peirastic or examination use of questioning, would be in an expert advice type of dialog. The questioner needs to test out the expert's assertion by seeing if it is consistent with other things the ex-

pert has said, or by seeing if it is consistent with what other experts say, and so forth. In such a context of use, Questioning Principle 1 could be justified. But its justification would be different from that for its use in a persuasion type of dialog, as outlined above. However, now having ventured into the unexplored territory of peirastic dialog, it is hard to pin down exactly what form this justification should take. But it does not matter, for present purposes. The point is that Hamblin has indicated, with great insight and foresight, how Questioning Principle 1, like Statement Principle 1, needs to be justified in different ways in different types of dialog, depending on the function that the speech act of asserting or questioning has in that type of dialog.

The problem with trying to make sense of the rules and principles discussed above is that the level of abstraction is too high. How to decide on which rules are right, or whether a rule should be formulated in some particular way, depends not only on the type of dialog, but on a lot of other things as well. The literature on informal fallacies has brought this sort of discussion about abstract rules of dialog down to earth by going into details on what sort of problems arise when you try to apply such a rule to actual cases of realistic argumentation. To put the discussion of the rules into a more practical setting, what needs to be done is to look at some of these problematic cases right in some shorter span of dialog where questioning, replying and asserting are taking place. There has been a tool developed in the new dialectic that is actually quite useful for this purpose.

11   The future of dialectic as a subject

Going through the whole history of dialectic, from its ancient Greek roots to its current revival, there are several key questions about how wide or narrow the subject should be taken to be. The Greeks took dialectic to be linguistic in nature – for them it was the art of conversation. But it was a special art of conversation that was different from rhetoric. Dialectic was seen as (what we nowadays call) a branch of logic by Aristotle. It was seen as a subject that used logical argumentation to question a respondent to uncover logical difficulties and problems, like self-contradictions, in the respondent's replies. Dialectic turned into a kind of technical subject in the Middle Ages. But it was later seen as trivial or useless. This negative view of dialectic was expressed by Kant, who equated it with idle chatter and illusion, seeing it as no different from sophistic. Others, like Hegel, wanted to make dialectic more than just an art of conversation, and to turn it into something beyond a linguistic subject. Nowadays the term 'dialectic' is most commonly associated with Marxist dialectical materialism, and thus it is doubtful to many whether dialectic has anything to do with logic. And yet, following Hamblin's lead, 'dialectic' is the

term now used for formal systems representing conversational argumentation. Modern logic has turned dialectic into a variety of formal dialectical systems representing different kinds of conversational exchanges in which two parties reason with each other. In modern logic, dialectical systems have been constructed that appear to represent persuasion dialog, information-seeking dialog, and inquiry. In logic, there appear to be no formal dialectical systems of deliberation, negotiation, or eristic dialog, although argumentation in these types of dialog has been widely discussed in many fields.

Although the Greeks were concerned especially with philosophical argumentation as a type of dialog identified with dialectic, their approach can actually be seen as being somewhat broader. As noted in chapter 1 and above, Aristotle, in a seminal but perplexing passage in *On Sophistical Refutations* (165a38–165b11) distinguished four kinds of arguments used in discussion. Didactic arguments are those that reason from principles appropriate to each branch of learning, and not just from the respondent's replies. Dialectical arguments, as noted above, are those that reason from generally accepted opinions to a contradiction. Examination arguments are used to test the knowledge of the respondent. Contentious arguments are those that only appear, but do not really reason from generally accepted opinions. This division of types of arguments does not square with the classification of types of dialog in the new dialectic. But it does show that Aristotle was concerned with arguments as used in different types of dialog, indicating a multiple view of argumentation that is consistent with the multi-context approach of the new dialectic. On both views, a given argument used in a particular case can be evaluated quite differently, depending on what type of dialog it was supposed to be a part of.

Given these historical developments, and the renewal of dialectic in modern logic, what is the best way to view it as a subject that is potentially useful? The best way is to think of it as an extension of the work of Grice (1975), that is, concerned with collaborative conversations in which moves, like arguments, are seen as contributions made at the appropriate stage of the conversation. Although Hamblin began to systematize the logical structure of argument moves in such a conversational framework, neither Grice nor Hamblin were in a position to classify the different types of goal-directed conversations in a systematic way. Now that the new dialectic has proposed such a classification, summarized in Table 1.2, it has become more evident how dialectic can be applied and why it is a useful subject. So conceived, dialectic is an art of rational conversation that begins to look a lot, in broad outline, like the Greek idea of it.

The best approach for the new dialectic is a revival of the Greek conception. On this view, dialectic is a branch of logic that is used to analyze and evaluate arguments (and related argumentative moves, like the asking of questions) in a context of conversational use. How the different types of conversations (dialogs) should be

precisely defined and classified is still subject to ongoing investigation. Dialectic is a structural subject and mathematical models, represented as formal dialectical systems, are central to its development. How dialectic will turn out as a subject depends on how the various formal dialectical systems are applied to real cases of problematic argumentation. Many of these are problematic borderline cases. For example, in some cases the same argument can start out as being in one type of dialog and then shift to another type of dialog, as will be shown in examples presented in chapter 7. As the new dialectic matures, it will become clearer how the various formal systems fit together and can be classified as representing distinctive types of conversational frameworks of argumentation.

For the present, what is vital in regaining support for dialectic as a respectable branch of logic is to assign to the term 'dialectic' a clear provisional meaning that defines it as a recognizable subject area suitable for exploration and research. The seven important characteristics of dialectic, as outlined above, provide a basis for developing this clear meaning. Like all proposed definitions advocated in academic pursuits at the beginning of an investigation, this one should be seen as a hypothesis that is put forward for acceptance, but is not carved in stone. Defining dialectic as a subject is itself a sort of dialectical activity. The new proposed definition is meant to move discussion of the subject forward, and to move the subject itself forward as well.

Persuasion dialog

Recent work in artificial intelligence has developed formal models of argumentation in persuasion dialog where the aim of one party is to rationally persuade the other party to come to accept a proposition as true (Bench-Capon, 2003; Prakken, 2006). Atkinson, Bench-Capon and McBurney (2004) describe the persuasion dialog as a type of dialog in which there are conflicting points of view and each of the two parties attempts to persuade the other of his or her viewpoint using rational argumentation in order to resolve the conflict. The one that has the more convincing argument, taking into account burden of proof, is the one who should be able to successfully persuade the other to come to accept his or her viewpoint. Given that successful fulfillment of the goal of persuasion is that of acceptance by the other party, however, it has been suggested that this kind of model of argumentation incorporates a kind of relativism by subordinating a truth requirement to an acceptability requirement. Since persuasion dialog has an adversarial aspect, it may seem that the winner is the arguer who presented his or her argumentation more effectively in winning an audience over, even though the argumentation of the other side might really be closer to the truth. The question is whether a successful persuasion dialog can be taken to move toward the truth of the matter being discussed or provide evidence that the proposition at issue in the dialog has been shown to be true or false, other than by merely getting a particular audience or respondent to accept it.

Boger (2005) argued that models of persuasion dialog in argumentation favor acceptance over truth, and thus lead to a pernicious kind of relativism. In this chapter it is shown that even though (a) it is often impossible to prove conclusively in persuasion dialog that one's conclusion is true or that one's opponent's conclusion is false, and (b) successful argumentation is based on acceptance of the other party, it does not follow that acceptance is being favored over truth. Based on computational systems of persuasion dialog in artificial intelligence using defeasible argumentation, it is shown how evidence of two kinds can be used to make an assessment of whether a persuasion dialog is successful: the maieutic depth of the dialog, and the standard of evidence met. Using this evidence, it can be judged

whether argumentation in a persuasion dialog has moved towards the truth of the matter being discussed or not.

The argument of this chapter is that the arguments advanced by both parties in a persuasion dialog can be tested and deepened by the probing criticisms opposed arguments presented by the other side. Although the argumentation put forward by a side is merely plausible to begin with, and it may even be defeated or shown to be fallacious as the dialog proceeds, the dialog may be successful if the argumentation is deepened in a certain way. Six factors are presented as the evidential basis that should be used to judge the depth of such a persuasion dialog. This evidence can then be used to make an assessment of whether the persuasion dialog has moved towards the truth of the matter being discussed or not. It is shown how new computational systems of argumentation developed in AI have thrown light on the nature of defeasible reasoning by analyzing the notions of how one argument can attack and defeat another argument, and how the utilization of proof standards in a persuasion dialog can be used to judge whether a given argument is acceptable or not as an investigation moves forward. It is shown how a proposition can be taken to be acceptable if the appropriate kind of evidence supporting it has been put forward in a dialog presenting a strong enough argument to overcome doubt and meet the proof standard for that type of dialog. It is argued that in persuasion dialogs on controversial subjects, where the need is to resolve a conflict of opinions about values (Bench-Capon, 2002), the task of the arguer is not to prove that a claim put forward is conclusively true beyond all doubt. It is argued that the success or failure of a persuasion dialog needs to be evaluated by standards of evidence used to show when a burden of proof has been met. Finally, it is argued that the idea that anyone has an exclusive claim to know the truth in a persuasion dialog is a hallmark of closed-mindedness that is antithetical to getting closer to the truth.

1   Persuasion in rhetoric and dialectic

The renewal of dialectic has led to much recent concern about the status and relationship of the two fields of dialectic and rhetoric. It has become evident that the new dialectic is useful for rhetoric – and for communication studies generally – as well as for logic and cognitive science (van Eemeren and Grootendorst, 1984; van Eemeren, Grootendorst, et al., 1996). But on the other hand, it does not seem to be a part of rhetoric. Rather, its origins and methods seem to fall within logic. On the other hand, the advent of the new dialectic seems to suggest that logic, or practical logic of the kind associated with dialectic, is much more closely connected to rhetoric than we thought in the past. Of course, there are the leading exceptions. Aristotle saw rhetoric and dialectic as closely connected. Because Aristotle has had

such a huge influence on the development of both subjects, and because even now his views permeate recent thinking on the subjects, to get a fresh approach it is necessary to re-examine the fundamental Aristotelian notions at the basis of dialectic and rhetoric. The rationale for this is by no means purely historical, or to try to grasp what Aristotle really meant. Rather the intent is to try to deal with certain preconceptions about rhetorical and dialectical argumentation deriving from prejudices about Aristotle's views. The prejudices are so firmly in place that it is hard to even question them. But it is only through analyzing them that some clearing of the ground can take place. Once that ground is cleared, quite a different view of the relationship between rhetoric and dialectic will emerge. However, it must be admitted that centering the analysis on Aristotle has real dangers. One danger is that of getting caught up in technical disputes among classicists and Greek philosophy specialists. An even more real danger is that of using Aristotle in a biased way to support one's own view of rhetoric or dialectic. Leff (1993, p. 314) has shown how contemporary scholars in rhetoric have often used Aristotle's *Rhetoric* as supporting "different visions of what rhetoric is or should be." Gross (2000, p. 26) has also shown how interpretations of Aristotle by recent scholars in rhetoric are used by these scholars to develop their own theories within an Aristotelian framework. Thus there is a danger of the tail wagging the dog. The danger is one of reinterpreting Aristotle to fit our current conceptions of what we think Aristotle should be saying. Despite these very real dangers, the peculiar history of rhetoric and dialectic makes it difficult to grasp what their relationship should be.

The concern of rhetoric, at first sight, seems to be quite different from that of logic. In the minds of many, rhetoric is associated with the persuasive use of arguments, with the function of persuasion, or with the study of persuasive discourse. The aim is evidently to assist an arguer to make a case for a conclusion by putting forward arguments that will support that conclusion in the sense of getting the target audience to accept that conclusion (when presumably, they did not accept it before the argument was presented to them). Rhetoric has often been concerned with the style of presenting a speech, in a way that makes it seem to have little to do with logic. Thus rhetoric and logic have been sharply contrasted. Logic has to do with the structural correctness of arguments, according to normative standards of reasoning. The aim of logic is to evaluate arguments as structurally correct or not. Rhetoric has to do with the effectiveness of arguments as used to persuade or influence a target audience. An argument could be structurally incorrect, but an audience could find it quite persuasive. On the other hand, an argument could be structurally correct, but an audience could find it quite unpersuasive. It seems that these two subjects are entirely different, and couldn't be farther apart. One has to do with psychological effectiveness of arguments and persuasion. The other is an

abstract subject that has to with the structure of arguments – their logical forms and how these forms are instantiated or not in particular cases of argument use.

But scholars working in the field of rhetoric generally don't seem comfortable with the idea that their goal is to produce arguments that are psychologically effective in persuading a target audience, no matter whether the arguments are logically correct or fallacious. On the other hand, they often act as though effectiveness is the main, or even the exclusive aim of their craft. This ambivalence has been noted by Schiappa (1995) and Jacobs (2000). Schiappa observed that despite the better impulses of the rhetorical theorists, what they actually do in practice in their work often seems to reduce to the issue of effectiveness. Jacobs (2000, p. 273) agrees that rhetorical analysts, while they don't accept this statement of their aim in theory, they "tend to accept it in practice". It can be argued then that even though many practitioners of rhetoric accept the view that effectiveness is the goal of their field, there remains room to question whether this assumption should be accepted uncritically. It could be that structural correctness is tied closely in with psychological effectiveness. And if so, ignoring structural correctness of argumentation, and viewing effectiveness as not tied to structural correctness, could be a way of practicing rhetoric that makes it less of an effective field. In short, this perceived difference between logic and rhetoric, when viewed too sharply or exclusively, may only reflect current practices. It could also be that these current practices need to be questioned and changed, in order to improve both fields.

Leff (2000) sees the differences between rhetoric and dialectic as more differences of degree, emphasis or orientation rather than as absolute differences that make the two fields into categorical opposites. Leff (p. 247) cites four such differences. The first is that dialectic deals with abstract issues while rhetoric deals with specific issues. The second is that dialectic deals with propositions and inferences, while rhetoric deals with how propositions relate to social norms and circumstances. The third is that dialectic proceeds by question and answer whereas rhetoric proceeds through uninterrupted discourse. The fourth is that dialectic uses technical language while rhetoric accommodates and embellishes language. It is not too hard to see that there are exceptions to all these general points of contrast, and Leff suggests (p. 247) that the tradition of rhetoric and dialectic is not as simple as these four differences suggest. As a corrective, Leff (p. 247) issues a caution not to exaggerate the differences between dialectic and rhetoric. For one thing, dialectic is more interactive and flexible than formal logic, does deal with real cases of persuasion, and derives its premises from social beliefs (*endoxa*). Thus it would appear that dialectic does have an empirical aspect, or at least that it is related to social norms and popular opinions that are accepted by an arguer or audience.

A key feature of the new techniques of using argumentation schemes in the new dialectic and the new rhetoric is the taking into account of the context of an argu-

ment. Newer styles of rhetorical analysis of argumentative texts (Zarefsy, 1990) use techniques of argumentation analysis to reconstruct argument structures in extended texts of discourse like political speeches. To conduct such an analysis, it is important to look at the purpose and context of a speech, as well as the individual arguments contained in it. This use of purpose and context puts rhetorical methods quite close, apparently, to the kinds of techniques of argument analysis used in the new dialectic. The new dialectic takes an applied point of view of evaluating arguments as used in a conversational context, following the conversational logic of Grice (1975). The structure of the individual argument is important. But what is also important is how that argument was used for some purpose in a context of discourse. These new trends seem to bring logic and rhetoric much closer together, making each of them closely related to the fundamental pragmatic Gricean notion of viewing moves of argumentation in relation to a collaborative talk exchange between two parties. Giving such a pragmatic notion of argumentation a place of fundamental importance makes for a revolutionary approach to both subjects. Logic is seen as much more practical and contextual in nature, while rhetoric is seen to be a more normative subject in which goals and conversational rules are important.

A new way of viewing rhetoric as a pragmatic subject has been proposed by Dascal and Gross (1999). They propose a marriage between Gricean pragmatics and classical rhetoric, based on the Aristotelian framework of rhetoric in (Perelman and Olbrechts-Tyteca, 1969) and the Gricean pragmatic theory of conversational interaction. These two frameworks fit together well, as Dascal and Gross have showed, because the forms of argument (argumentation schemes) studied by Pereleman and Olbrects-Tyteca can be quite well modeled in the Gricean conversational setting. Dascal and Gross show, for example (pp. 121–125), how Aristotle's theory of rhetorical persuasion based on *pathos* and *ethos*, can be modeled as forms of argumentation giving rise to conversational implicatures within the Gricean framework of a collaborative talk exchange. What their analysis also brings out, however, is the pragmatic structure of forms of argumentation like *ad misericordiam* and *ad hominem* that, even though they were traditionally regarded as fallacies, have been recognized in the new dialectic as extremely common forms of argument that can often be quite reasonable in conversational argumentation. However, they are reasonable by the standards of the new dialectic, when viewed as used for some collaborative purpose with a speech partner within a goal-directed conversational framework. What is shown is that the judgments in a given case of whether an argument is fallacious, or is a reasonable instance of an argumentation scheme, depends on the kind of conversation the argument is supposed to be part of. An argument, thus conceived, needs to be seen as a contribution made at some stage of a goal-directed conversation.

Gricean pragmatics is shown by Dascal and Gross to be a new foundation for rhetoric that has outstanding potential to give rhetoric a new power to deal with argumentation in a way it never could before. But at the same time, the Gricean conversational analysis of implicature gives logic a new potential as an applied subject that has the power of analyzing and evaluating arguments, and other moves in argumentation, like the asking of questions, in a context of use in a given case. Many traditional problems of applied logic, like distinguishing between explanations and arguments, and finding missing assumptions in an argument, now become possible to solve. Grice's conversational pragmatics, by itself, was not adequate to provide the formal structures and methods needed to solve these problems. But the new dialectic, built on the Gricean conversational logic, can provide such methods.

Rhetoric is primarily a practical art. Its primary goal is to provide techniques that are actually useful to persuade an audience to accept a conclusion, or carry out an action. Dialectic is an abstract and normative subject, but its primary goal is also practical. Its primary goal as a useful subject is to take actual cases of argumentation and analyze and evaluate that argumentation critically, finding the stronger and weaker arguments in it. It goal is to provide methods that enable a person to judge which arguments are stronger and which are weaker. It is good enough for rhetoric to persuade an audience by using arguments that the audience thinks are strong, never mind whether they are really strong or not, from a dialectical or logical point of view. In fact, if there is any conflict between winning over the target audience and using an argument that is logically and dialectically correct and free from fallacies, rhetoric should always go on the side of finding the right argumentation to win over the audience. Why then should rhetoric have any use at all for dialectic? The answer resides in the following subtle point. An audience will be persuaded by arguments it thinks are strong, and these are generally arguments that are either strong, or that bear a resemblance to those that are strong. For this reason, the new rhetorician can actually learn quite a bit from the new dialectic.

What has been made clear is that there is quite a close connection between the new rhetoric and the new dialectic, and yet the two subjects are quite different, both in aims and methods. They both make use of the same kinds of argument structures, even though they make use of them in different ways. They both are based on the central notion of persuasion, but rhetoric is more concerned with effective persuasion that can change the beliefs of an audience or lead them to different actions, while dialectic is centrally concerned with rational persuasion. Dialectic is a branch of logic. It could be called practical logic or applied logic, but is often now (somewhat inappropriately) called informal logic. Rhetoric is often seen as a purely empirical subject, but once its cognitive component comes to be based on the new dialectic, it will become a much more interesting and powerful subject than it has ever been before.

2   Characteristics of persuasion dialog

The terms 'persuasion dialog' and 'critical discussion' are sometimes used interchangeably, as noted in chapter 1, section 5, but the critical discussion has been classified as a special type of persuasion dialog (Walton, 1989). The term 'persuasion dialog' has now become a technical term of argumentation technology in artificial intelligence (Bench-Capon, 2002, 2003; Greenwood, Bench-Capon and McBurney, 2003; Prakken, 2006), and thus a careful definition of it is required. As Bench-Capon (2002) has shown, in cases of disagreement about values[1], it is impossible to provide conclusive demonstration that proves a proposition beyond doubt. The role of argumentation in such cases, as Bench-Capon sees it, is to persuade rather than to prove, recognizing that the strength of an argument will vary according to its audience, and the comparative weight that the audience gives to the social values the argument advances. A value is a reason that is given to support a goal. If an arguer has the value of respect for the truth of a matter being discussed, it will support her goal of persuading the other party in a discussion by using rational argumentation. The value of truth is not in possessing it, or thinking that you possess it. Only the dogmatic or fanatical arguer thinks that she possesses the truth, and that all she needs to do is to persuade the other party to accept it. Argumentation on matters of values is inherently defeasible, and when arguing about such matters one needs to be open-minded. Thus even though it may be difficult to define the truth, and to know for sure when one has it, one can engage in rational argumentation in persuasion dialog with a respect for the truth, in a discussion that leads toward the truth as an ideal.

As shown by Bench-Capon (2003) and Greenwood, Bench-Capon and McBurney (2003), formal dialectical systems can provide a rational basis for the acceptance or rejection of such arguments. The system of Dung (1995) is often taken as a starting point, because it is minimal, in that it evaluates arguments only in how they are attacked by other arguments. A set of arguments can be visualized as nodes in a graph, and arrows lead from some nodes to others indicating which arguments attack other arguments. Using these attack relations, criteria can be set out to determine which arguments are acceptable and which are not. For example an argument that is not attacked is acceptable. If an argument has attackers, it can be defended against them by other arguments that attack them, and if so it is acceptable. Dung's system has been extended by Bench-Capon (2003) to allow for

---

1.   Here we will make no attempt to determine which particular disagreements, or types of disagreements, are about values, or partially involve values. Legal arguments are clearly about values as well as facts, but it can be questioned whether argumentation in scientific investigations is also partly about values. Certainly many of the most common kinds of persuasion dialogs are about values as well as factual issues.

evaluation of arguments as acceptable or not based on values. Given an ordering on values, the system allows a determination to be made of whether an argument is acceptable or not depending on the values expressed in it.

Other dialectical systems are based on standards for acceptance appropriate to the type of dialog disputants are engaged in. A range of such standards has been proposed, including ones used in law like preponderance of evidence and beyond a reasonable doubt (Farley and Freeman, 1995). These standards are now widely used in formal dialectical models of argumentation (Gordon, 1995). Prakken (2006) provided a review of existing formal systems of persuasion dialog. In a persuasion dialog, according to his account, two or more participants aim to resolve a difference of opinion, each trying to persuade the other to come to accept his or her point of view. As shown by Prakken, such systems for persuasion dialog have now found applications in various fields, including computer science, artificial intelligence, law, multi-agent systems, nonmonotonic logic, intelligent tutoring, and computer-supported collaborative argumentation. Prakken has shown how the rules for persuasion dialog can be contrasted with those for other types of dialog widely studied in the literature on argumentation and artificial intelligence, including negotiation, information-seeking dialog, deliberation, inquiry, and the quarrel, the so-called eristic type of dialog that is a verbal substitute for a fight.

In formal models of persuasion dialog there are only two parties, called a proponent and a respondent.[2] The dialog is adversarial. Each party has the goal of presenting arguments that are stronger than the other side, as well as at the same time attacking the arguments of the other side in order to weaken them. Because of this adversarial aspect, it may seem that the persuasion dialog is nothing more than a quarrel, and that the most vocal or aggressive party wins. However, according to the formal models of persuasion dialog, although each party has this competitive incentive to win, and therefore to construct as persuasive arguments as possible, the other side also has the capability of criticizing those arguments by finding weaknesses and fallacies in them. Also, the model of the persuasion dialog is rule-governed, and arguments are not acceptable as rationally persuasive unless they meet structural requirements.

'Persuasion' in this sense refers not to psychological persuasion but to an instance where a proponent gets a respondent to accept a proposition by presenting an argument that has a valid form (that of an argumentation scheme) and premises that the respondent is already committed to. Persuasion, in this technical and normative sense of the term, refers to the change in the respondent's commitments (Walton and Krabbe, 1995). The simplest way to explain the basic structure of a

---

2.   However, persuasion in a debate or a trial is more complex, because the outcome is decided by a third party who evaluates the arguments put forward by each side.

persuasion dialog is to use the diagram of (Walton, 1989, p. 6) shown in modified form in figure 3.1.
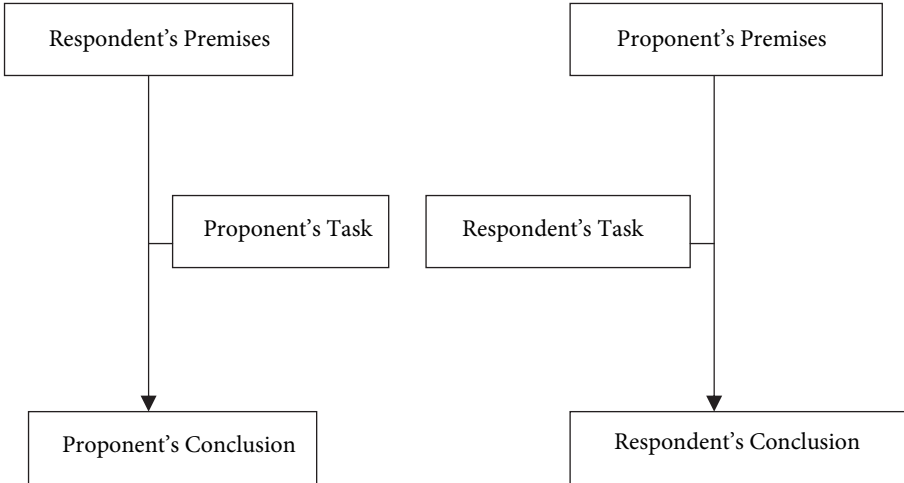

Figure 3.1 The structure of argumentation in a persuasion dialog

Each party has a central task to perform in the dialog, and the successful carrying out of that task by using rational argumentation is the goal of each party. The proponent's task is to rationally persuade the respondent to come to accept her (the proponent's) thesis. The respondent's task is to rationally persuade the proponent to come to accept his (the respondent's) thesis. Although figure 3.1 explains the essence of persuasion dialog in its simplest terms, there are various other factors that also need to be considered.

A successful act of persuasion consists of the proponent's presenting a structurally correct argument (by the appropriate standards) where the respondent accepts all the premises, or where the proponent has shown by his past moves in the dialog that he is committed to them. Assuming the standards for structural correctness are binding on the respondent as a rational arguer, he becomes committed to the conclusion (unless he can undercut or defeat it). If the proponent can carry out this task, she wins the argument, so to speak. She has successfully persuaded the respondent to accept her thesis. The thesis is a proposition (statement) assigned to a party at the initial stage of the dialog where the parties agree to use rational arguments to carry out their assigned tasks. So far, so good, but now we must define successful persuasion in a more complex way that allows for the possibility that the proponent cannot carry out her task in one step. She may have to

use a lengthy chain of arguments to persuade the respondent one step at time. Accordingly, we now introduce the following definition of 'successful persuasion'.

> The proponent persuades the respondent to accept a designated proposition *A* as true if and only if the proponent puts forward a chain of argumentation meeting the following requirements. First, each step, or single inference in the chain, is a structurally correct argument according to some appropriate requirements set out in the opening stage of the dialog. Second, the premises of the argument are all propositions that are already commitments of the respondent in the dialog. Third, the ultimate conclusion of the chain of argumentation, at the final step of inference, is the proposition *A*.

To grasp this definition, one has to come to understand that many of the arguments in the sequence used to persuade are defeasible, meaning they may be reasonably accepted at one point in the dialog, but then later retracted as new evidence enters into consideration. To accommodate this kind of argumentation, it needs to be possible for the respondent to sometimes retract commitment to a proposition that was formerly accepted. This feature models the important characteristic of rational argumentation in persuasion dialog that an arguer needs to be open to new evidence, and to the possibility of defeat. She cannot not just stick dogmatically to her opinion, once that opinion has been refuted by undercutting or defeating argument that meet success standards appropriate for the dialog. On the other hand, a respondent should not always be free to retract, in any situation, for example, if it starts to look like she might lose.

If free retraction at any time were to be allowed in a persuasion dialog, the proponent could never carry out the task of successfully persuading the respondent to accept her thesis. There must be some "bite" to rational argumentation, provided it meets the appropriate standards. Table 3.1 shows what happens in a dialog when the proponent puts forward an argument with premises and a conclusion, and the argument fits the form of one of the argumentation schemes.

Table 3.1 Bite of argument in a persuasion dialog

Committed to premises	Committed to conclusion	Argument not necessary
Committed to premises	Uncommitted to conclusion	Must commit to conclusion or ask critical question
Uncommitted to premises	Committed to conclusion	Argument not necessary
Uncommitted to premises	Uncommitted to conclusion	Can question a premise

If the respondent was already committed to the conclusion, the argument was not necessary, and there is no effect on the respondent's commitments in the dialog. If the respondent was uncommitted to the conclusion, but also uncommitted to the

premises, the argument would not force him to change his commitments in any way. For he can always merely reply by indicating he is not committed to one of the premises. It is only in the remaining row of the Table 3.1 that there is a significant outcome on the respondent's options. Either he asks one of the appropriate critical questions matching the argumentation scheme, or he must change his commitment set by becoming committed to the conclusion. He can then go on asking critical questions or posing counter-arguments, but if he runs out of them, he must at that point commit to the conclusion.[3]

The communal goal of the critical discussion is to resolve the conflict of opinions, but the individual goal of each participant is to win the argument by fulfilling its burden of proof, thus defeating the other party. Thus on the surface it looks like the goal of the critical discussion is not that of reaching the truth. The perception therefore is that the critical discussion is all about reaching agreement, and there is no regard for the real truth of the matter being discussed. This view is superficial however. Although the goal of the critical discussion is to get at the truth of the matter being discussed, typically on issues discussed in a persuasion dialog, like the philosophical issues of abortion or the issue of whether God exists, or the kind of issue resolved by a legal trial, it is unrealistic to think that the participants might actually arrive at the truth on the issue by the concluding stage of their discussion (Bench-Capon, 2002). Typically in such a case, the most successful kind of discussion we can have is one in which both parties probe into the reasons given by the other side supporting its viewpoint. In such a case, we can often say that the critical discussion was highly successful because the reasons behind each viewpoint were articulated and brought forward during the argumentation, and the weak points in each were pinpointed (Greenwood, Bench-Capon and McBurney, 2003). This kind of discussion could be extremely successful, even though there might be no way to know beyond doubt that the truth of the matter was reached. Still, we say in such a case that, at the concluding stage, it was shown that the argumentation put forward by the one side for its thesis was stronger than that of the other (Bench-Capon, 2003). The argumentation can be evaluated in a formal structure that combines each of the single arguments, chaining them together, and assessing whether the chain of reasoning leads to evidential support of an ultimate conclusion (a proposition at issue in the dialog) that meets the standard for that type of dialog. On a preponderance of the evidence standard, for example, this conclusion can be taken as proved, because it is stronger as evidence than the argumentation put forward by the opposed side.

---

3. Note that this rule does not apply to complex contexts of persuasion dialog like that of legal argumentation. Nor will it apply unreservedly in real instances of conversational argumentation, except where they are held to be bound by the normative model of persuasion dialog.

According to Bench-Capon (2003), it is impossible to demonstrate, in many real cases of disagreement, that either party is wrong. On this assumption, the role of argumentation in such cases should be to rationally persuade rather than to prove a conclusion beyond doubt. A persuasion dialog can be successful even if it does not resolve the issue being discussed conclusively, by proving that one side is right and the other is wrong. On the Bench-Capon model of persuasion dialog, whether an argument is successful or not depends on the comparative strength of the values advanced by that argument. To model this view of persuasion dialog, a value-based argumentation framework (VAF) is used. The VAF can be applied to explaining how a successful persuasion dialog can move towards the truth of a matter. The central goal of a persuasion dialog is to resolve the conflict of opinions, if that is possible, but if that is not possible, the aim is to at least bring the argumentation on the issue closer to the truth. This is done by examining the evidential roots of the arguments on both sides in a probing and critical way. Achieving the truth is not the goal of the persuasion dialog, or at least the principal one, for as Bench-Capon noted, in many instances, especially where values are involved, achieving such a lofty goal is not possible. However, respect for the truth should be a value underlying the argumentation, if a persuasion dialog is to reach its central goal. The value of bringing the evidential roots of the arguments on both sides to light in a persuasion dialog is that it reveals the truth of the matter being discussed. Thus a value underlying a persuasion dialog as a whole is that of respect for the truth.

How does a persuasion dialog move towards uncovering the truth of the matter being discussed even if it fails to resolve the issue decisively by proving that one side's thesis is true while that of the other is false? The answer has to do with the depth of the discussion. Consider a typical example of a philosophical discussion on some topic like the meaning of truth, or whether euthanasia should be allowed. It is unrealistic to expect that such a discussion will resolve the issue conclusively one way or the other by proving beyond all doubt that the thesis of one side is true and that of the other side is false. Still, the discussion could be quite enlightening and successful if both sides used strong arguments and criticisms to interact argumentatively with the position of the other side, and if, through the discussion, the position of each side was refined and expressed with more precision and clarity. There are two benefits to such a discussion. One is the refinement of one's own view, making it not only more sophisticated, but based on better reasons supporting it. The other is the increased capability to understand and appreciate the opponent's point of view. For it is typical of a deep and thoughtful discussion of this sort that it leads a participant to a better understanding of the arguments on the other side, while at the same time it leads to a sophistication and refinement of one's own view, by making it able to respond to the objections posed by the other side.

Classical examples are the Socratic dialogs of Plato. True, they're a bit one sided, and artificially contrived, perhaps to defend Plato's own views, or to illustrate the argumentation skills of Socrates. But they are genuinely edifying reading, and very instructive, just because of the depth of the dialogs. They take perennial issues like truth and justice, and examine some of the most common and widely held arguments on both sides. Even though the conflict is not resolved decisively one way or the other, in most instances, the discussion itself is very beneficial to readers because of the argumentative interaction displayed, and the reasons supporting each of the opposed views put forward in a lively way. Many philosophers after Plato have preferred the dialog method, in some form or other. They include Aquinas, who used the philosophical method of putting forward not only his own views, but the leading objections to them along with the replies to these objections. This methodology gives depth to a discussion by bringing out the arguments on both sides.

3   Defeasibility and acceptance

Most formal systems of argument in persuasion dialog used in artificial intelligence are nonmonotonic, meaning that a set of premises is taken to imply a conclusion in a sense meaning that the conclusion is taken to be a defeasible consequence of the premises (Prakken, 2006, p. 174). In this section it is shown how defeasibility is an important characteristic of arguments and how artificial intelligence has drawn important distinctions between ways one argument can attack and defeat another. Pollock (1995) drew a widely known distinction between two kinds of arguments that can defeat another argument: rebutting defeaters and undercutting defeaters. A rebutting defeater gives a reason for denying a claim (Pollock, 1995, 40). Thus a rebutting defeater attacks the claim, or conclusion of the argument it is aimed at. An undercutting defeater, in contrast, attacks the connection between the claim and the reason rather than attacking the claim directly (p. 41). It is often thought to be puzzling how this latter notion works exactly, and so it is best to outline the standard example of it often cited.

Pollock's example of the red light (1995, p. 41) is a very good illustration of how an undercutting defeater attacks an argument. In this example, initial appearances proved an argument giving ground for acceptance of a conclusion, but then new evidence comes in that defeats the original argument by showing it no longer can be held to support the conclusion that was originally accepted.

> For instance, suppose x looks red to me, but I know that x is illuminated by red lights and red lights can make objects look red when they are not. Knowing this defeats the prima facie reason, but it is not a reason for thinking that x is *not* red.

After all, red objects look red in red light too. This is an *undercutting defeater* (Pollock's italics in both instances).

The sequence of reasoning in Pollock's example can be analyzed below, showing that it is based on generalizations that link observational premises to conclusions in two stages.

First Argument

Fact: This object looks red to me.

Generalization: When an object looks red, then (normally, but subject to exceptions) it is red.

Conclusion: This object is red.

Second Argument

Fact: This object is illuminated by a red light.

Generalization: when an object is illuminated by a red light this can make it look red even though it is not.

Undercutting: Withdrawal of the prior conclusion that this object is red.

The second argument undercuts the original argument because it attacks the connection between the claim and the reason. On Pollock's account, (p. 41) the argument is classified as an undercutting defeater, but not a rebutting defeater, because of the second argument, based on the generalization that red objects look red in red light too.

New argumentation systems for artificial intelligence are being devised that aim to model defeasible reasoning that can be used under conditions of uncertainty and lack of knowledge of the kind that would enable a discussion or investigation to prove a conclusion is true or false beyond all doubt. Because much of the reasoning used in the fields of AI like robotics and multi-agent systems is of this defeasible kind, in which evidence can rationally support or undercut a claim, but cannot prove or disprove it beyond all further argumentation or investigation, such systems are very useful.

Carneades is a formalized argumentation framework, named after the ancient Greek skeptical philosopher.[4] The purpose of Carneades is to provide a computational model of argument structure and evaluation that can examine arguments

---

4.   Carneades was the head of the third Platonic Academy that flourished in the second century B.C. He was born around 213 B.C. in Cyrene or Cyrenaica (now in Libya), and died around 128 B.C. Because he did not write anything, and the notes of his lectures taken by his student Clitomachus do not survive, it has often been assumed that Carneades was not a significant philosopher. But according to Hankinson (1995, p. 94), he not only had an unparalleled reputation in his lifetime as a master dialectician, but is "one of the great figures in the history of philosophy".

pro and con an issue being discussed, and resolve the issue by evaluating the argumentation in relation to proof standards. The proof standards define whether an argument is acceptable or not in the context of the discussion or investigation that is going forward in which arguments are continually being updated by new evidence that comes in. The primary motivation of Carneades is to model critical questions of a kind used to attack an argument of a particular type. Typically, these arguments tend to be defeasible, and therefore need to be evaluated on a basis of how well the proponent of the argument responds to critical questions asked by the respondent in the dialog. In Carneades, two kinds of critical questions are distinguished. When the respondent asks one kind of question the proponent's original argument is defeated unless he gives an appropriate answer right away. For example, suppose that the proponent puts forward an argument from expert opinion and the respondent asks for some evidence the expert has based her opinion on. Failure to answer such a question defeats the argument. When the respondent asks the other kind of question, he needs to support it with further evidence before it will make the proponent's argument fail. For example, suppose the respondent asks whether the expert is personally reliable. Is she honest? The proponent can reply, "Prove that she's not, if you think she isn't". In Carneades, additional premises to the ordinary premises in an argumentation scheme are classified as assumptions or exceptions. An assumption is taken to hold, unless it can be disproved, whereas an exception is taken not to hold unless evidence to support it can be produced. In Carneades, critical questions matching a scheme can be represented as additional premises in the scheme itself, either as an assumption or an exception. Carneades is a computational model, in the sense that its mathematical structure consists of a set of computable functions. The model has been implemented and tested using standard programming languages. In chapter 5, section 3, it will be shown how posing critical questions can be represented in a dialog system.

It is interesting to see how Carneades models Pollock's concept of an undercutting defeater. Argument structures are represented in Carneades by directed graphs (figure 3.2) consisting of argument nodes and statement nodes (Gordon and Walton, 2006, p. 7).

As shown by the pair of boxes in the middle of the diagram, the first argument, called a1, has two premises. One premise is that the object in question looks red. The other is the general rule, that things that look red are red, applies to the case at issue. Thus at the top part of the diagram we have the two premises supporting argument a1, which support the conclusion that the object is red, shown in the box at the top. The closed arrowhead leading from the applicable premise to argument a1, is shown as a assumption. The ordinary line leading from the premise 'looks red' to the argument a1 has no arrowhead, indicating it is an ordinary premise. Now we turn to the bottom part of the diagram, where new evidence is intro-

duced, to the effect that in this particular case, the object is being illuminated by red light. This new information constitutes an exception to the general rule that things that look red are red, because the things under consideration are illuminated by a red light, showing them as red whether they are really red or not. So once the new evidence of the red light, indicated in the bottom box in the diagram, is introduced, it provides a new argument that works as an exception to the rule that things that look red are red. In this special kind of situation, the general rule is defeated. Thus when we look over the argument diagram as a whole, we see that the original argument with its two premises no longer gives sufficient evidence to support the conclusion that the object is red. Carneades the philosopher would have recognized Pollock's red light example, because it is a perfect example of an argument based on appearances, that is reasonable, but not beyond skeptical doubt. It can either be supported or refuted by new evidence that has entered into consideration as an investigation proceeds.
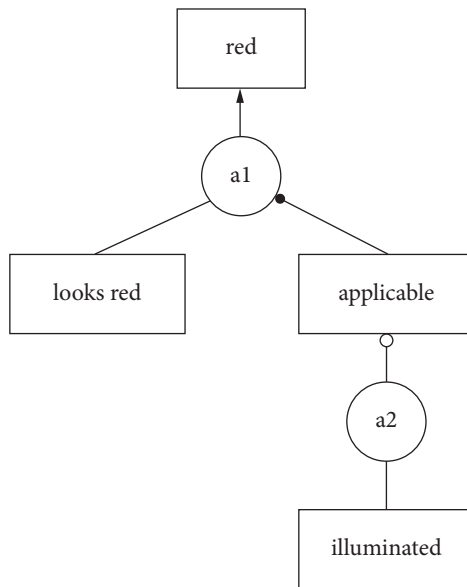

Figure 3.2 Pollock's red light example in carneades

Carneades developed his criterion of rational acceptance based on what seems to be rationally acceptable according to three criteria. The first criterion is that the thing to be accepted, a proposition or something presented as true, actually does appear to be true. When a subject experiences a "presentation" (something that

appears to him), he can accept it tentatively as true if it seems convincingly to be true (*AL*, 168–170). This criterion models defeasible reasoning, for the presentation in question can be accepted as true, even though it is not known for sure to be true, and might later be shown to be false. The second criterion is that a presentation should be accepted if it is plausible in the first sense and is also irreversible, meaning that it should fit in with other presentations that also appear true (*AL* 176). The third criterion is that of the testing of the presentation (*AL* 182–183). Sextus Empiricus cites the famous case of the snake-rope to illustrate the use of this third criterion along with the previous two criteria (*AL* 188). A man sees a coil of rope in a dimly lit room. It looks like a snake to him. He accepts the proposition that it is a snake. Basing his prudent action on what he has accepted, he jumps over the object, keeping his distance from it. But then he turns back. He sees it did not appear to move at all, even though he had jumped right over it. These appearances provide a reason for accepting the proposition that the object is not a snake. The initial appearance that the object is a snake has turned out to be reversible, thus violating the second criterion. The man then retracts his prior commitment, and now reasons that the object was probably (plausibly) not a snake. He then moves to the plausible assumption that it might have been a rope. This hypothesis seems to better fit the existing appearances. But he could still be wrong. Snakes are sometime motionless. Hence the man carries out a test. He prods the object with a stick. If it moves, then it is a snake. If it fails to move even when prodded, it is a rope. Of course, neither conclusion is absolutely certain.

Carneades' example of the snake-rope illustrates very well how defeasible reasoning works in modern AI argumentation systems, especially in the Carneades AI system, specifically designed to model his approach. The Carneades system is designed to support conversational exchanges that enable argumentation to be tested as new data comes in, and implicit premises to be revealed dynamically as a dialog proceeds. Carneades does not require all premises of an argument to be made explicit when the argument is put forward by a proponent. Instead, during critical questioning, implicit premises of the argument are revealed. Indeed the critical questions themselves can be modeled as implicit premises of an argument. The structure of the argument is one factor in how it is evaluated, whether its premises are accepted is another. But a third factor is how well the argument stands up to critical questioning in a dialog that tests it by subjecting it to probing examination and attack by an opposing arguer.

The third feature of Carneades that is important here is that the system is based on acceptability or defensibility of a statement, not truth or falsity. Three proof standards are defined: scintilla of evidence, preponderance of evidence and dialectical validity. These proof standards are ordered, scintilla of evidence being the weakest, and dialectical validity being the strongest. The acceptability of a

statement depends on its proof standard, and whether a statement satisfies a proof standard depends on the defensibility of the arguments pro and con that statement. The defensibility of an argument depends not only on whether its premises hold, but also on the structure of a particular type of argument, and the original status of the statement in the dialog, whether it is an issue, undisputed, accepted or rejected. Thus Carneades provides a recursive testing procedure that can be applied to argumentation of the kind typically found in a persuasion dialog. In this kind of argumentation, there is generally insufficient evidence to prove that a statement is true, or to prove beyond all doubt that it is false, but there is enough evidence for a defensible argument for the conclusion that the statement is acceptable because it meets a proof standard.

One of the interesting things about Carneades is that it enables critical questions to be anticipated within the structure of the argumentation scheme itself. For example suppose a proponent puts forward an argument from expert opinion. The scheme for this type of argument tells us that there are three ordinary premises. One is that the source cited is an expert, the second is that the expert has mastery of knowledge in some specific domain, and the third is that the expert has made statement S. The conclusion is that S should be acceptable to the respondent as a commitment. But there are additional premises in the argumentation scheme that strengthen the argument by dealing with possible critical questions before the respondent can raise them. These implicit premises have what can be described as a procataleptic function in the argument. 'Procatalepsis' is a traditional term used to refer to a figure of speech in which the speaker raises an objection to his own argument and answers it within the argument himself. Hence the procataleptic function is the strengthening of the proponent's argument by dealing with criticisms, objections or counter-arguments at one move in a dialog that might be advanced by the respondent at some later move.

4   Evidence, testing, and burden of proof

The way that any discussion or investigation moves forward towards proving or disproving a claim at issue is through a process of generating hypotheses that are then tested and refined. But this testing process works in different ways in different types of dialog. In a critical discussion, testing takes place through probing criticisms made by one side against the arguments put forward by the other side. In a scientific investigation, an initial hypothesis is formed, based on the data or given appearances. The hypothesis is an attempt to explain the data, often by fitting it into a theory, or even by inventing a new theory. The hypothesis is then tested by collecting new data, and seeing how well the hypothesis can explain the new data

in addition to the previous data. The hypothesis may pass or fail the test. If it fails the test it may have to be refined and reformulated. Thus a new hypothesis is formed that, once again explains the existing data, but may have to be given up as new data comes in. This process explains the defeasibility of scientific reasoning, fitting with Popper's philosophy of falsification. On such a philosophy, summarized at the end of this chapter, a conclusion can be accepted if it has been shown to be based on evidence, even though the argument supporting it is inherently defeasible. On this view, even scientific argumentation is subject to defeat as further evidence come into an investigation. Such evidence leads to rational acceptance, but not proof of truth.

What is evidence? Basically, the answer to this question is provided by the following definition. A set of propositions can be classified as evidence for another proposition if and only if the inferential relation between the initial set and the designated conclusion proposition takes the form of an argument that supports the conclusion. Such an argument needs to be evaluated as evidence by three criteria. One is that the inferential link between the premises and the conclusion need to conform to some structure that represents an argumentation scheme. Another is that the premises have to be verified, supported or established to such an extent that they are deemed acceptable. The third is that the argument has to be strong enough to rebut opposed arguments called refutations. What is evidence can be decided by Carneades' three criteria. An agent has evidence that a proposition is true if (1) the proposition appears to be true, (2) it fits in with other propositions that appear to be true and are accepted as true, and (3) it is tested, and passes the test. Evidence that a proposition is true may be weak if only the first criterion is met, but becomes stronger if the second is also met, and even stronger if all three are met. To the extent that the evidence is stronger supporting one hypothesis, and the evidence supporting each of the competing hypotheses is comparatively weaker, the one hypothesis is closer to the truth than the others.

Van Eemeren and Houtlosser (2002) see the task of assigning burden of proof as a matter of procedural agreement, seeing it as varying with the type of dialog and with the type of procedural agreements necessary for the success of that dialog. However they do formulate one general principle: "Unless it is explicitly agreed otherwise, the burden of proof is on the side of those whose standpoints are challenged by the other party"(p. 18). They concede that there can be a transfer of argumentative duties from the one party to the other in some cases, but they describe such a transfer as a "shift of initiative" rather than a shift of burden of proof. Under their analysis, such shifts of initiative are commonly accompanied by strategic maneuvering on the part of both parties to a discussion. Such strategic maneuvering about the burden of proof has in the past been studied in connection with fallacies. For example it has been observed that some fallacies are based on

subtle shifts in the burden of proof that can be powerful argumentation tactics, but on the other hand can sometimes be subtly deceptive.

Another approach to the relationship of burden of proof to defeasibility is the argument-based approach, based on the observation that defeasible forms of argument, by nature, leave room for counter-arguments that can be brought to bear on them. Logics of defeasible argumentation to formalize common-sense reasoning have been developed in artificial intelligence, especially as applied to legal reasoning (Gordon, 1995; Prakken and Sartor, 1996, 1998; Bench-Capon, 2003; Verheij, 2005). These systems are especially designed to capture arguments that can hold tentatively, but can later be defeated due to the adversarial nature of legal argumentation. According to Verheij (2003), a form of argument called DMP (defeasible *modus ponens*) can be formalized in his system for defeasible logic called DefLog as follows. The first premise of the scheme is symbolically represented using a connective for defeasible implication, ~>.

> $A \sim> B$
>
> A
>
> Therefore B

The first premise could be paraphrased as stating, 'If *A* is true then generally, but subject to exceptions, *B* is true'. According to Verheij (2001, p. 232), DMP is a contingent argumentation scheme, meaning that it can depend on the particular circumstances whether it can be used to form an inconclusive argument that holds.

Verheij's system has been specifically applied to defeasible argumentation using a tree structure that can contain arguments having the form of strict *modus ponens* (SMP) or DMP at each node of the tree. The defeasible arguments of the DMP type allow for counter-arguments attacking the given argument by arguing that there is an exception to the qualified (defeasible) rule (Prakken, 2006, p. 5).

Argument-based logics of this kind presuppose a dialectical structure in which argumentation needs to be analyzed and evaluated on two levels. Argument-based logics accomplish this task by modeling argumentation as a tree superimposed on another tree (Prakken, 2007, p.6). First there is a tree in which propositions are linked to each other by inference rules. Second, the dialectical status of arguments is tracked in a secondary tree that maps all the possible ways to support or attack a given argument. This secondary dialectical tree tracks connected sequences of moves and countermoves in a dialog between a proponent and a respondent. Just as in the case of default logics, this secondary level represents the tracking of a chain of argumentation through an investigation as arguments are modified in the face of new evidence and conclusions are supported and accepted, or later defeated and retracted. Such a tree models what Prakken (2007, p. 6) calls reinstatement,

where one argument attacks a second one defeating it, but then later, a third argument comes in supporting the first argument by defeating the second one. Thus the first argument is said to be reinstated. Reinstatement raises problems of retraction of commitment of the kind studied in (Walton and Krabbe, 1995), suggesting that tracking of commitments in a dialog is involved.

When argumentation is modeled in this way, a given argument can be extrapolated forward to consider how it might be questioned, attacked or defeated by future moves of the respondent in the dialog. For example, in Carneades, the critical questions matching an argumentation scheme can be expressed as premises that are supplementary to the ordinary premises of the schemes, and that have the procataleptic function of taking an opponent's possible future critically challenging moves against the argument into account. This way of viewing defeasible argumentation has many implications for dialog theory that have not so far been fully recognized, but that will be explored in chapter 7. It implies, for one thing, that some central notions of argumentation theory, like the speech acts of assertion, assumption and presumption, and the notion of burden of proof, need to be rethought. As well, the connecting of the notion of defeasible argumentation with dialog theory brings with it a need to rethink the concept of evidence that links evidence with burden of proof.

Asserting a proposition always characteristically carries with it a requirement of burden of proof. If I assert the proposition that snow is white, for example, and my assertion is questioned or challenged by someone who doubts it, I am obliged to justify the proposition by offering an argument or justification for it, meaning that I must offer a reason for accepting the proposition as true. The speech acts of presumption and assumption are different. I can assume that any proposition is true without cost, meaning that there is no burden of proof attached. I can assume a proposition is true even if I know it is false, or others think it is false. Presumption is halfway in between. I can presume a proposition is true as long as this proposition is not known to be false. However, once it has been shown by strong enough evidence that this proposition is false, I must at that point give it up as a presumption. In other words, moving forward with the presumption is based on a kind of open-mindedness, meaning that the proposition presumed to be true might default in the future, and at that point the person who put it forward might have to retract it. A hypothesis is a proposition that has not been proved, but has the status of a presumption that has some evidence in favor of its truth, and can be provisionally accepted provided it has not been proved to be false.

5   Dialogs, truth and relativism

Some kind of relativism is inherent in the argumentation approach, to the extent that this approach is a systematic effort to analyze and evaluate arguments not only at a semantic level, but also at a pragmatic level. The aim is to analyze and evaluate how arguments (and other speech acts) are used for some conversational purpose in a context of dialog. From the semantic perspective, an argument is seen as a set of propositions with an arbitrarily designated one called the conclusion. From the pragmatic perspective, an argument requires a dialog setting in which one party doubts a particular proposition while the other party presents some other propositions brought forward to remove this doubt. The pragmatic approach does seem to involve some kind of relativism, as it makes the evidential worth of an argument dependent on the framework of dialog in which it was put forward. But is it an acceptable relativism or a bad relativism that infects the informal logic movement and the recent research work on argumentation like a virus?

Many in traditional logic may be hesitant to accept informal logic (argumentation theory) because of worries about relativism. The same kinds of worries motivate suspicions that logicians often have about rhetoric, that it has no concern for truth, and is merely concerned with persuading an audience. Argumentation theory attempts to analyze and evaluate how arguments are used in everyday conversations and other contexts. Many might say that even though it is a good model of this kind of argumentation, it has not yet been established how it is related to the finding of truth. If so, it is an open question whether it is based on a relativism that subordinates acceptability to truth. Some kind of relativism might be expected in any model that analyzes how argumentation is used in realistic contexts of conversation. But a relativism that subordinates acceptance to truth seems to be a bad kind of relativism.

The charge of relativism has often been worrisome, even to those in the informal logic movement. Van Eemeren and Grootendorst (1995 p. 124) expressed doubts about Perelman's notion of the universal audience on grounds that it provides only a weak criterion for argument evaluation that may be open to a worrisome kind of relativism. Tindale (1999) and Freeman (2005) have defended acceptability as a criterion for resolving disputes, whereas Siegel and Biro (1997) and Johnson (2000) have argued for retaining a truth criterion, recognizing that some form of relativism might be a dangerous result of using an acceptance criterion for evaluating argumentation.

An issue to be decided is whether agents that search for the truth can actually grasp the truth, or come to know it, or whether truth is something objective that is intrinsically external to our attempts as agents or researchers to reach it. Supporting the latter view is the opinion of many scientific researchers that the results of their research are best seen as only approximations to the truth. Indeed, it would

seem that some version of Popper's principle of falsification (Popper, 1963) would require that scientific findings should always be regarded as defeasible, meaning that should new evidence come in, it may require modification or refining of the original hypothesis, or even giving it up altogether in favor of a new hypothesis. In sympathy with his view, there has been considerable work recently on the notion of verisimilitude or approximate truth in science. In this view, truth is never really obtainable in this world, because of limitations on knowing what is true or false beyond the possibility of doubt. Still, on this view, truth can function as a goal or underlying value that our investigations in research can be aimed at. So conceived, truth functions as an ideal that can be reached, but we can never show, in any real investigation, that it has been reached.

Even if persuasion dialog may help us move further toward the truth of the matter being discussed, it may not be realistically possible to know whether the critical discussion has actually arrived at the truth at its concluding stage. However, it may be that we can judge whether the conclusion of such a process is closer to the truth or further away from the truth. This idea of closeness to the truth has motivated recent discussions in the philosophy of science about what is called truthlikeness or verisimilitude. As Graham Oddie (2001) pointed out in a recent survey article on truthlikeness, this notion has been a latecomer to the philosophical scene.

> Truthlikeness is a relative latecomer to the philosophical scene largely because it wasn't until the latter half of the twentieth century that mainstream philosophers gave up on the Cartesian goal of infallible knowledge. The idea that we are quite possibily, even probably, mistaken in our most cherished beliefs, that they might well be just *false*, was mostly considered tantamount to capitulation to the skeptic. By the middle of the twentieth century, however, it was clear that natural science postulated a very odd world behind the phenomena, one rather remote from our everyday experience, one which renders many of our commonsense beliefs, as well as previous scientific theories, strictly speaking, false. Further, the increasingly rapid turnover of scientific theories suggested that, far from being established as certain, they are ever vulnerable to refutation, and typically are eventually refuted, to be replaced by some new theory. Taking the dismal view, the history of inquiry is a history of theories shown to be false, replaced by other theories awaiting their turn at the guillotine.

Sir Karl Popper (1963) was the first philosopher to make a serious attempt to use the notion of truthlikeness in the philosophy of science. It is well known that Popper argued for a kind of asymmetry between falsifiability and verifiability of scientific theories. He argued that scientific theories cannot be conclusively verified, and should always be seen as open to defeat in the future, as new data comes in and scientific theories are refined. Thus he argued that the only kind of progress an

inquiry can make is through falsification of existing hypotheses and theories. However, he added that some false hypotheses can be closer to the truth and others (Miller, 1974). Thus he recognized the notion of verisimilitude or truthlikeness, and many other philosophers of science since then have taken up this idea (Kuipers, 1987). There appears to be little agreement on how the notion should precisely be defined, but the basic idea of verisimilitude is widely accepted as an important concept in current philosophy of science. But how could it be applied to the critical discussion, something not widely treated in the literature on the philosophy of science, except by some like Popper and Kuhn who have made some use of the notion.

6   The charge of pernicious relativism

Boger (2005, p. 187) warned that there is a "specter of relativism" that haunts argumentation theory, even calling it a "pernicious relativism" (p. 187). He calls it "a tragic flaw" that has led to some traditional logicians to forecast the demise of informal logic and argumentation theory (p. 188). Boger (p. 188) cites the enthusiastic remark of Ralph Johnson that "this multiplicity of currents, rather than being a deficiency of the movement, speaks to its strength, its integrity, and the promise of further accomplishments". While agreeing with Johnson about the usefulness of informal logic and argumentation studies as opposed to the more traditional approach of teaching logic, Boger (p. 189) levels the serious criticism that the "pernicious relativism" of this newer approach has a tendency to "subvert objective knowledge". It is worth quoting his charge, to appreciate the seriousness of it.

> We have … become troubled by a noticeable tendency within the argumentation movement toward psychologism and a related relativism that might subvert an aim to secure objective of knowledge. We believe that this tendency is a consequence of subordinating what has been a central truth requirement of a good argument, replacing it with an acceptability requirement, and strongly relating argument appraisal to audience adherence and context.

Note carefully the nature of the charge. Boger does not explicitly claim that the argumentation movement has ignored truth altogether, or advocates the kind of relativism that says that one argument is as good as another, or that there is no rationality. His charge is that the argumentation movement has subordinated what he calls a central truth requirement of a good argument to an acceptability requirement. Nor does he claim that the argumentation movement evaluates arguments exclusively in relation to audience acceptability or adherence. Rather he

claims that the argumentation movement strongly relates argument evaluation to these factors.

The last claim is perhaps not too surprising, given that the argumentation movement has stressed the importance of contextual factors for evaluating real examples of argumentation in a given text of discourse. The claim of psychologism is not very worrisome or credible either, as argumentation theorists from Hamblin (1970) onwards have been very careful to distinguish between commitment and belief (Walton and Krabbe, 1995). It is the charge of subordinating truth to acceptability that is more worrisome, because it attributes a lack of concern for the truth to the argumentation approach, apparently leaving it open to the worries about pernicious relativism and subverting objective knowledge. On Boger's account, the cause of these worrisome tendencies is the subordination of truth to acceptability, and in fact Boger is right that the mainstream of argumentation theory hasn't said very much about the traditional notions of knowledge and truth that have been central to philosophy, especially in the field of epistemology. Certainly it is right that these matters deserve further attention.

This kind of relativism might be harmless, if the dialog context can be modeled using a formal dialog structure, and if an argument is evaluated using inference rules of deductive logic, or other argumentation schemes that have a precise structure. Hamblin (1970), and following Hamblin, Walton and Krabbe (1995), have adopted the concept of commitment in dialog as an important ingredient of their criteria for argument evaluation in formal dialog systems. But this move may seem open to a charge of relativism. The criteria of a successful argument in a system of formal persuasion dialog are that the argument should be structurally correct, meaning that it is valid or fits argumentation schemes, and that the premises should be commitments of the other party in the dialog. These criteria allow for some measure of objectivity, for such an argument has to fit an argument form, like a form of argument in deductive logic or an argumentation scheme, and it also has to fit formal requirements concerning how it was used as a move in a dialog system. Still, the argument does not have to be based on true premises. Instead, it only needs to be based on the commitments of the opposing party. These criteria may be open to criticisms of relativism, because successful persuasion is defined in terms of commitment without the need for meeting any truth requirement.

The persuasion dialog is not the only context of argumentation use postulated in argumentation theory and artificial intelligence (Prakken, 2005). Other types of dialog that contain argumentation have been recognized. There can be different standards of proof in different contexts of argumentation, research, investigation and inquiry. Each different context for collecting and evaluating argumentation has its own standard (Gordon, 1995). What can be said about how argumentation in each of these other types of dialog moves toward or away from the truth? We

cannot offer complete answers to these general questions in this paper, but still, some discussion of them is called for.

The goal of the deliberation dialog is for the participants to decide what to do in a situation in which there is uncertainty and lack of knowledge. Deliberation can be solitary, but it has a dialog format because the deliberating agent needs to find the strongest arguments on both sides of the issue. Deliberation can also take place among groups of agents – for example, in a town hall meeting on whether to go ahead with a proposal to install a new sewer system. One of the main types of arguments they must use is called argumentation from consequences. According to this form of argument, if the proposed course of action is found to have positive consequences, that is a reason for adopting it, while if it is found to have negative consequences, that is a reason for not adopting it. Truth doesn't have much of a place in a deliberation dialog. The goal of deliberation is not to find the truth of the matter being discussed, but rather to arrive at a decision on a course of action to take, given particular circumstances. Truth can come into it indirectly, because an intelligent deliberation needs to be based on information, and not just any information, but information that can be shown to be reliable or true. Nevertheless, the central goal of the deliberation is that of choosing a reasonable course of action among the available alternatives.

Negotiation is a type of dialog in which there is a conflict of interests, and the participants engage in argumentation in order to reach some agreement on how to resolve it. In persuasion dialog truth matters a lot, because each participant is trying to persuade the other to come to accept a proposition as true or false. The goal is different in negotiation dialog from what it is in persuasion dialog. In negotiation, matters of truth and falsity of propositions don't matter so centrally, because what matters is the interests of each side. Interests can sometimes be measured in financial terms, but factors like prestige or perceived status can also a subject of negotiations in some cases. In a principles negotiation, matters of truth may come into the process indirectly, but they are not central. In negotiation each side makes offers and concessions in a process of bargaining in which truth is not the goal the process is aiming at.

In the information-seeking type of dialog one party lacks information and the other tries to provide it. For example, suppose I am a tourist in a foreign city and I approach a person and ask her, "Where is the Central Station?" I want reliable information, but the cost of being wrong is not that high. I am not trying to prove something beyond doubt, but just trying to find out something as information I can use. Often there is a shift from an information-seeking dialog to a persuasion dialog. For example, if we are having a persuasion dialog on the issue of euthanasia, it will be much better informed and will be a much deeper discussion of the issue, if it is based on relevant medical facts. Or to cite another kind of example,

suppose a legislative assembly is having a debate on whether it would be a good idea to install a new dam. Such a discussion would have to be based on many facts about ecology, about the cost of the project, and about the impact it would have on the population living in the area. Thus for such a persuasion dialog to have depth, it is necessary for it to be based on facts as well as values, or at least on information that all parties to the discussion can accept as factual.

The inquiry is different from the information-seeking dialog because a high standard of proof is required to reach the concluding stage. A group of investigators collect all the relevant data and only then reach the argumentation stage. The goal of the inquiry is to prove whether some designated proposition is true or false, or if that cannot be achieved, to prove that it cannot be proved whether it is true or false. The burden of proof that marks a successful concluding stage of such an inquiry is very high. The inquiry only accepts as premises propositions that can be verified, so that there will be no need in the future to have to retract any of them. The goal of the inquiry is to prove that a particular proposition is true or false by verifying or falsifying it by objective evidence. The goal of the inquiry is to provide such a solid proof so that the need to have to make retractions in the future is eliminated. The method used is to draw inferences very carefully and to accept as premises only propositions that can be established as true or false.

One view of scientific research called foundationalism postulates that scientific investigation should take the form of an inquiry. Aristotle and Descartes, both working scientists as well as philosophers, advocated this view. Descartes even adopted the view that premises in an inquiry should be established as true, beyond doubt. These views have largely been discredited by contemporary philosophers. In practice, the goal of proving a scientific finding beyond doubt is unobtainable in scientific research, because of the principle of falsifiability. In fact, scientists do sometimes have to retract previous findings, because of new data or the formulation of a better theory. Nevertheless the goal is to take considerable care to eliminate the need for retraction, and therefore the standards of what is accepted as verified in a scientific inquiry are meant to be high. Here we would argue that they are too high to represent any real scientific investigation. We will not adopt the view that a scientific investigation takes the form an inquiry of a kind that is cumulative, i.e. one in which retraction of commitment is never allowed. We will opt for a weaker model in which an investigation moves forward by collecting data, formulating hypotheses to explain the data, and then testing them. We will use the term 'investigation' in this weaker sense, but hold that the methods of evaluating argumentation in an investigation of the scientific type are different from those properly used in a critical discussion. Many of the kinds of arguments are the same, but the standards for evaluating them and the methods of testing them are different.

Eristic dialog is a quarrelsome type of dialog in which two parties "hit out" at each other, in which one or both sides often use *ad hominem* arguments claiming that the other side's argument must be wrong because he is an unethical person who is therefore dishonest and untrustworthy. The goal of each participant is to articulate some hidden or unarticulated grudge or grievance, and by bringing it to light, attack the other party personally. Eristic dialog is associated with the committing of fallacies and the use of deceptive tactics. It is not meant to discover the truth of a matter, and more often leads away from the truth by distracting an opponent or an audience. The participants in a quarrel may claim to have truth on their side, and each may attack the other party as being untruthful, but the aim is not really that of finding the truth.

7   Judging the maieutic depth of a persuasion dialog

The extent to which a persuasion dialog is successful in moving towards its goal of reaching the truth of the matter being discussed can be judged by the depth of the discussion. Determination of depth is a separate matter from determination of which side won the discussion by meeting its burden of proof. The depth of the discussion can be decided by the participants or perhaps even better by a third party, a referee or judge, who examines the whole critical discussion as it took place, and arrives at a determination of how thoroughly each side probed critically into the weaknesses of the arguments of the other side. A factor important in determinations of the depth of a critical discussion is that of deciding whether the strongest relevant arguments on both sides were brought forward. To make such a judgment, the argumentation on each side needs to be represented by an argument diagram, a tree structure showing the premises and conclusions in the sequence of argumentation in the given discourse.

The issue of defining depth in a persuasion dialog is linked to the completeness problem, the problem of when a respondent should be obliged to stop asking critical questions in response to an argument, and either accept the conclusion of the argument or provide some good reason for not accepting it. The problem is whether there should be some sort of procedural rule that puts an end to such a sequence of critical questioning. This difficult problem has resisted solution so far, but is a central one related to key notions of argumentation theory like burden of proof and argument sufficiency. One approach to solving the problem put forward here is to utilize the notion of depth of argumentation to define a limit of critical questioning once it has reached a designated depth. This approach links the notion of depth to the function of critical questions in argumentation. According to one theory of the nature of critical questions, the function of asking critical questions

is to probe into the weak points of an argument. This probing process throws light on the argument in several ways. It reveals the weaknesses of the argument to the respondent, so that he can see better how to attack it and raise doubts about it. At the same time by revealing these weaknesses, it can show the proponent how to strengthen the argument by filling in gaps and missing assumptions. By bringing out such gaps and missing assumptions, it can reveal deeper arguments on which the original argument was based, providing both parties, as well as observers, with insights about evidence on both sides of the issue being discussed.

When confronted with an argument based on an argumentation scheme, a respondent might be said to have three options. The first is the asking of appropriate critical questions. The second is the raising of objections. For example, Hamblin (1970) analyzed the fallacy of equivocation as a procedural objection that might be raised in a dialog in response to an argument put forward. The respondent might object that the proponent's argument contains a key ambiguity and is therefore open to the objection of equivocation. Such a response should be classified not as a critical question matching an argumentation scheme, nor as a counterargument, but as an objection. The third option is for the respondent to put forward a counterargument, meaning an argument that has the opposite conclusion of the original argument put forward by the proponent.

In light of these three options, a broad definition of the notion of depth in a persuasion dialog can be formulated, based on the following six questions.

(1)   How many critical questions have been answered?

(2)   Have all of them been answered?

(3)   How many objections have been replied to?

(4)   How many counterarguments have been put forward?

(5)   How many counterarguments have been rebutted by counter counter-arguments?

(6)   How lengthy is the chain of rebuttals and replies?

By counting up all the factors listed in the six questions, a numerical estimate of the depth of the argumentation in any given case can be given. This way of defining the depth of a given instance of argumentation, provides a measurement of depth useful for computational purposes, for example in artificial intelligence.

Some tools are available from artificial intelligence for modeling the notion of the depth of a search. Current search technologies are used to solve problems like how to get from one city to another. The search process is usually represented as a search tree that starts from an initial node and seeks for paths leading towards a goal node representing a solution to the problem (Russell and Norvig, 1995, p. 71).

One type of search strategy is the breadth first search in which the root node is expanded first, and then all the nodes generated by the root node are expanded, and then this process is repeated (p. 74). A depth first search begins the search at the deepest level of the tree, and then only when it hits a dead end does the search go back and expand nodes at shallower levels (p. 77). The basic principle in both types of searches is that one node is selected as the start point and from there its successor nodes are examined and explored. There is one kind of search called an adversarial search in which a game tree representing all possible moves by both players is searched to find the most effective playing strategy for each player.

The notion of depth of argumentation can easily be modeled in a Dung-style system. The more arguments a given argument is attacked by, or defended against arguments that attack it, the deeper that given argument is. This notion of depth could be extended to arguments based on values of the kind represented by the system of Bench-Capon (2003). This notion of depth of argumentation is only a very simple one, but it could be extended to model a fuller notion of depth useful for argumentation studies. This project has not yet been carried out, but what can be done here is to set a target for it by explaining how depth of argumentation in a persuasion dialog needs to be defined in a fuller fashion.

In judging the depth of a given instance of a persuasion dialog that has been carried to the concluding stage, a search has to be made to determine whether the strongest relevant arguments on both sides were brought forward, and how thoroughly each side probed critically into the weaknesses of the arguments of the other side. These two factors alone, however, are not sufficient for arriving in a determination of the depth of the discussion. The judgment needs to be based on the following six factors:

(1)   how many of arguments were brought forward defending the arguments on each side,

(2)   how many of these arguments were undercut or defeated,

(3)   how many implicit premises were revealed in these arguments,

(4)   how revealing these implicit premises were in producing a maeutic effect,

(5)   how well the discussion was informed of the relevant facts on the issue,

(6)   and how strongly the argumentation throughout the whole dialog, supported or refuted the fundamental thesis at issue.

These six factors are the evidential basis that should be used to judge the depth of a persuasion dialog. This judgment is a comparative rating. Comparing two texts of discourse each of which represents a persuasion dialog on some particular is-

sue, an examination of the argumentation used in each discourse, based on this evidence, can arrive at a decision of which had the greater depth.

A judgment of the depth of the persuasion dialog can be used as part of the evidence to make an assessment of whether the dialog as a whole moved away from the truth or towards the truth of the matter being discussed. If the dialog considered all the important arguments on both sides of the issue, not only putting them forward but also criticizing them in a deep manner, then the dialog may be judged to have moved towards the truth of the matter being discussed. This judgment means that the persuasion dialog threw light on the issue being discussed by examining and evaluating the strongest arguments on both sides. What is meant when it is said that this dialog moved towards the truth of the matter being discussed is that it took us beyond the point where we were before the discussion took place. It moved toward the truth if it threw light on the issue by probing into it more deeply, thereby enabling both the participants and the onlookers to come to a deeper understanding of what the issue is really about. This effect can be achieved only by pitting the strongest arguments on each side against each other in an adversarial contest that exposes the fallacies and critical shortcomings in the arguments of each side. A successful persuasion dialog of this type takes us closer toward the truth of the matter being discussed by bringing out evidence for and against the propositions on each side of the original conflict of opinions.

A hypothesis put forward at the discovery stage of a scientific investigation will not be as close to the truth as one that has been theoretically refined to a higher degree, and that has passed more tests comprising more data. The key notion is that of depth of search in an investigation. As more and more of the data have been searched more thoroughly, and as the hypothesis in question has been tested more rigorously, thus meeting a higher standard of burden of proof, it can be concluded that an investigation achieves greater depth. The notion of depth is used here in the same sense as that used in the analysis of how a critical discussion can come closer to the truth.

Now it has been shown how a critical discussion can be judged to have moved closer to the truth of the matter being discussed. To make such a determination, the evaluator has to look over the four stages of the dialog, especially the arguments and criticisms put forward by both sides during the argumentation stage. She then has to judge whether the discussion has moved toward or away from the truth of the matter being discussed by making a judgment of the depth of this argumentation.

As shown in figure 3.3 below, there are two parallel streams in a persuasion dialog. The type of dialog is shown in column 1, setting the requirements for the confrontation and opening stages. The argumentation stage is shown in column 2. Once the burden of proof is met by one side or the other, as shown in column 3, the argumentation can move to the concluding stage (column 4).
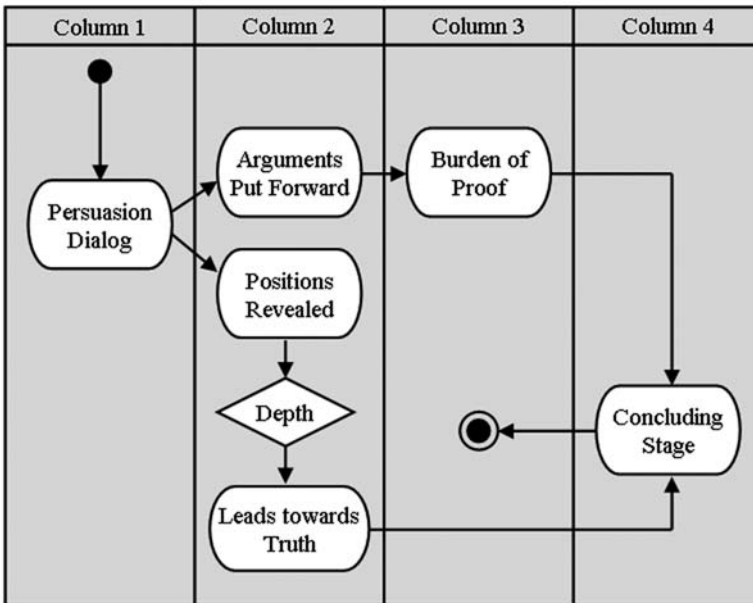
Column 1	Column 2	Column 3	Column 4

Figure 3.3 Parallel streams in a persuasion dialog

The argumentation in the persuasion dialog divides into two streams in column 2. The top stream is the acceptance stream, where arguments are put forward by both sides, based on the commitments of the other side and on argumentation schemes that are used to infer new commitments from previous ones. The bottom stream is the truth stream, where positions of both sides are revealed, and then, judging by their depth, a judgment can be made whether the argumentation has led toward the truth of the matter being discussed.

Generally then, even though the goal of argumentation in a persuasion dialog is to reach the truth of the matter being discussed, in many cases it is unrealistic to expect that this goal will be achieved even though the discussion could be quite successful. The best we can realistically hope for in such a discussion is that each side will probe critically into the weaknesses in the arguments of the other side, so that fallacies, contradictions, and other logical weaknesses are revealed. But achieving this lesser outcome can often be very valuable, and bring us closer to the truth of the matter being discussed, because both sides have to refine and articulate the views that they defend. Take a philosophical discussion in which two opposing viewpoints are put forward, and strong arguments are put forward supporting both viewpoints. There may be no clear winner or loser at the concluding stage, but even so, the argumentation in the dialog may have been enlightening, by having a maieutic effect of bringing new ideas to birth. This effect is achieved not by

finding the truth of the matter beyond all doubt, but by critically examining the errors, weaknesses and fallacies on both sides through a process of probing scrutiny into the evidential reasons behind both viewpoints.

8   Aiming at the truth

A central problem with the standard of evidence theory of truth, as with all theories of truth, is how one knows a proposition is true. This is the central problem of epistemology, and there's not enough space to attempt to meaningfully discuss it here, much less to attempt to solve the problem. However, it might suffice to say that any theory of truth or knowledge inevitably traces back to the fundamental insights of the Greek skeptics, who pointed out that we are often deceived, even about propositions we are strongly convinced are true. Thus the dogmatic notion that a proposition can be taken to be true without any reservations about whether we know it is true, in the cases of propositions accepted in science or propositions reporting immediate sense data, needs to be carefully reviewed before discussing any theory of truth or knowledge. The notion of truthlikeness is based on a certain skeptical background, meaning that a lot of the propositions we say we know to be true in everyday life can be subjected to skeptical doubts of various kinds. As noted by Graham Oddie, mainstream philosophers in the twentieth century gave up on the Cartesian goal of infallible knowledge (Oddie, 2001). On Carneades' view of reasoned argumentation and evidence, truth may be an ideal that is aimed at, but we cannot be sure in any serious investigation that we have reached it, beyond all possibility of doubt. To find a middle position between that of dogmatism and skepticism, Carneades argued that we can accept evidence based on appearances, provided it is consistent with other appearances, and is tested in a manner that presents additional evidence for its acceptance as true. While acceptance in such cases is defeasible, it can lead to stronger acceptance of a proposition through a process of testing. By this means we can judge whether an investigation has collected sufficient evidence to meet the standard of acceptance that is appropriate. By this means we can make an evaluation of the extent to which the investigation is moving forward by an accumulation of evidence indicating that progress is being made toward revealing the truth or falsity of the proposition being investigated.

The standard theories of truth that have been put forward by philosophers are taken to fall into two categories. According to the robust theories, truth is a property, in the way that being white is said to be a property of snow when the proposition 'Snow is white' is asserted. Robust theories strive to explain this property and define its central characteristics. There are five leading robust theories (Kirkham, 1992). The correspondence theory defines truth as correspondence with an objec-

tive reality, that is, a state of affairs in the world. On this theory, the proposition 'Snow is white' is true if and only if it expresses a state of affairs that corresponds with the objective reality that in fact snow is white. The coherence theory defines truth as a relationship of coherence of the designated proposition within some wider set of propositions. On this theory, coherence is thought to be stronger than consistency, and is usually defined in terms not only of consistency in the logical sense, but also in terms of the comprehensiveness of the whole set of propositions involved, as well as incorporating notions of justification and evidence. The consensus theory defines truth in terms of the set of propositions agreed upon by a group of agents. It is often held that argumentation theory is based on a consensus theory, because the concept of critical discussion that is central to argumentation theory appears to be fundamentally based on the underlying notion of agreement. The pragmatic theory of truth defines truth in terms of the practical consequences of a proposition. On this theory, to say a proposition is true essentially means that its acceptance has practical utility of some kind. The social constructivism theory of truth holds that truth is constructed by social processes, and is relative to cultural standards, or historical social and institutional norms. Argumentation theory may also seem to some to be based on a social constructivism view of truth, because a conflict of opinions is resolved in a critical discussion through an adversarial process in which a proponent and an opponent engage in a power struggle to see who has the stronger argument that can defeat the viewpoint of the other party, thus resolving the conflict.

Contrasted with robust theories of truth, recently there have been many more modest theories put forward by philosophers called deflationary theories. The most famous of these is the semantic theory of truth proposed by Tarski (1944, 1956). This theory is associated with Tarski's famous equivalence: for any sentence p, p is true if and only if p. This theory was developed for formal languages, like first order logic, and is based on the distinction between ground level and meta-level languages. The premise of the theory was that a language cannot contain its own truth predicate, and therefore in order to define truth, we have to move from the ground level language that Tarski called an object language to a metalanguage, or secondary language about the object language. The problem with the semantic theory is precisely that it is a deflationary theory, and doesn't tell us very much about what the essential (central) characteristics of truth are. It appears to only tell us that truth can be eliminated, and that saying a proposition is true is really nothing more than asserting it at a higher level.

Kripke (1975) put forward a theory of truth that is similar to the semantic theory, except that it argues that a natural language can contain its own truth predicates. His theory begins with a set of sentences in natural language, and defines truth for a subset that predicates truth or falsity just to the sentences in the subset.

This theory shares the deficiency with other semantic theories in that, while it does enable us to avoid contradictions and paradoxes of the kind encountered traditionally in first order deductive logic, mainly having to do with self-reference, it appears to tell us little about the central defining characteristics of the notion of truth. The problem with semantic theories is that they're not very helpful when it comes to explaining what truth is in the way that the robust theories attempted to do. It would seem to a reasonable expectation that a theory of truth should present enough of its central characteristics to explain what truth is, even if the explanation links truth to other problematic notions like justification or evidence.

This problem of failure to offer enough of an explanation to be of any use seems to attach to all deflationary theories that reject the idea that truth is a robust concept. What is common to many of the deflationary theories is that they follow Tarski's equivalence formula, but do not go far enough further to meet the explanatory requirement of a good definition or theory of truth. They stop short at merely drawing attention to the disappearance of the quotation marks as the defining characteristic of the notion of truth.

Deflationary theories that make any attempt to go beyond this limitation seem to reduce to forms of the consensus theory or the social constructivism theory. For example, on one theory, to say that snow is white is true is to perform the speech act of agreeing with the claim that snow is white. This theory is an interesting one, in that it links the notion of truth to speech acts, and that does seem like a helpful move in the direction of meeting the explanation requirement. But it fails because it does not go far enough into the details of the various kinds of speech acts that are related to truth, merely reverting back to the notion of agreement as the foundation on which to base the theory of truth. The problem with this kind of theory is that by using agreement as the basic notion, it simply collapses back into the consensus or social constructivist view, holding that truth is whatever is agreed upon by some group, or whatever is historically or culturally accepted.

To sum up, none of the philosophical theories of truth seem to be all that helpful in providing some definition of truth that would be useful to guide us in making judgments of whether or to what extent an argument moves toward or away from the truth. The deflationist theories lack enough explanatory power to be useful for this purpose. The traditional theories appear somewhat plausible, but are open to so many objections, and seem so unclear on what is meant by key terms like correspondence and coherence. Also, there appears to be no agreement on which theory is best. Truth may be the target or goal of the argumentation in critical discussion or scientific investigation, but exactly what is the argumentation supposed to be aiming at? To be aiming at the ultimate goal of possessing the truth, or of conclusively showing that the proposition advocated by the other party in a persuasion dialog is false, are unrealistic goals, judging from the skeptical

point of view of Carneades. Is there some other way we can frame persuasion dialog as being aimed towards the truth?

9   Truth, evidence and acceptance

It may be possible to move the discussion a little further forward (toward the truth, we might say), by presenting some remarks about the relationship of truth and rational acceptance that might be helpful. The first problem for any philosophical discussion of truth and evidence is to decide on the so-called truth bearer, the type of entity that is supposed to carry the properties of being true or false. On the account proposed here, propositions will be said to be the truth bearers, and for our purposes propositions will be treated as equivalent to statements, but different from sentences. We adopt the usual convention that propositions are expressed in sentences. In other words it is possible, on this approach, for two different sentences to express the same proposition. A proposition may be defined as what is asserted in an assertive sentence. An assertive sentence can be contrasted with other forms of speech, like a question or an imperative sentence, one that expresses a request or command. Assertion is defined as a form of speech act broadly following the analysis of it given by Searle (1969). As a type of speech act, assertion needs to be contrasted with the closely related speech acts of assumption and presumption. The distinction between them has to do with burden of proof. Assertions made in a dialog need to be proved by the party who ventured the assertion, but surely they do not always have to be proved beyond all doubt to be true to warrant acceptance of the proposition claimed. Demanding such a high requirement would block a typical persuasion dialog rather than moving it forward.

Truth is an ideal guiding the argumentation in an investigation or discussion by showing that the marshalling and assessment of evidence has moved toward the ideal by the concluding stage. Truth can be partially uncovered by an investigation that moves systematically through stages toward its concluding stage. Indeed, such an investigation is designed to move toward the ideal of truth. The extent to which the investigation has moved toward the truth is measured by judging the depth of the argumentation in it. This definition of truth requires that a distinction be drawn between subjective and objective standards representing evidence. Truth is an objective standard that is external to the investigation itself and its four stages. 'Objective' here means independent of the agent who holds that the proposition in question is true or false. 'Subjective' means dependent on the knowledge of the agents in the investigation. Hence this approach to truth is based on some prior notion of the definition of an agent (Wooldridge, 2002). Only once the notion of an agent is defined can the distinction between subjective and objective standards

representing evidence be clarified. The agents in the investigation cannot possess truth, or at least be sure that they do. The best they can do is to aim at truth, and conform to the ideal of truth by following a procedure aimed at the truth.

The view of truth and acceptance based on this definition could be called the standard of evidence and depth (SE&D) model. The agents are the arguers who participate in a dialog, like a persuasion dialog, and their arguments can be evaluated as successful or not in achieving the goal of the dialog according to two factors: (a) whether the argumentation at the closing stage meets the standard of evidence appropriate for that type of dialog, and (b) the depth that the discussion has achieved by the closing stage. In figure 3.4, it is assumed that the standard of evidence for the type of dialog is identified as the beginning of the sequence. But here we are only concerned with applying the SE&D model to persuasion dialog. Only after the concluding stage, once the depth of the discussion has been evaluated, can a judgment be made on whether the argumentation has moved toward revealing the truth of the matter being discussed or not.
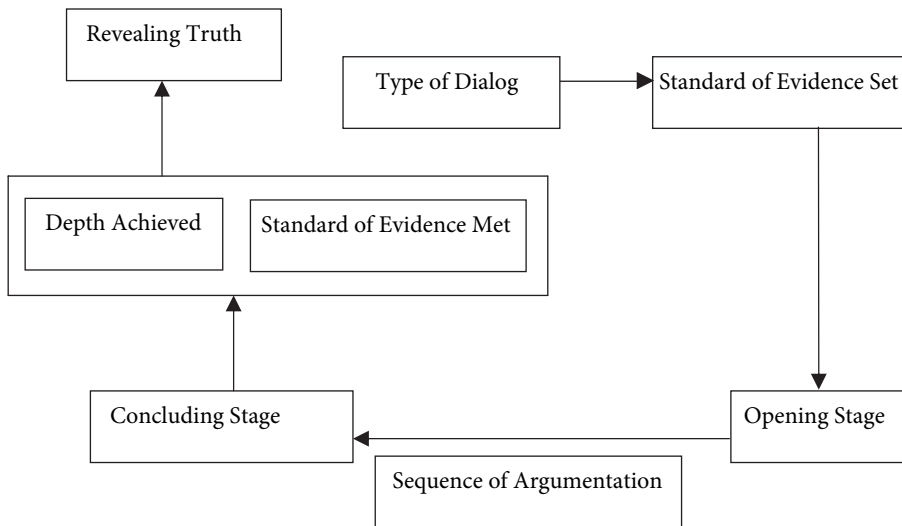
Figure 3.4 The SE&D model of argumentation moving toward truth

On the SE&D model, truth is an objective ideal that can be used to evaluate the worth of argumentation that has reached the closing stage in a dialog and has been evaluated by standards representing evidence. In addition to persuasion dialog, the SE&D model can also be applied to the kind of scientific investigation that collects evidence and forms hypotheses on the basis of that evidence. How far a dialog or investigation has moved toward the truth can be judged by an assessment of the

depth of argumentation reached by the dialog, or the depth of searching for and uncovering of relevant evidence in the investigation. The notion of the depth of argumentation in a persuasion dialog was defined in section 7, but different ways of measuring depth may have to be applied to different cases. But what are the standards for evidence in an investigation?

One standard of evidence is that provided by appearances (empirical data). Another is provided by specifying normative standards, for example by a system of definitions, axioms, logical rules and theorems in a scientific theory. A proposition that conforms to standards set for a body of other propositions that systematically fit together is true. Thus both the correspondence and the coherence theories are partly right. In some instances, a proposition is judged as true because it corresponds or conforms to empirical input (appearances). In other instances, a proposition is judged to be true because it is coherent with other propositions that form a system of definitions, axioms, logical rules and theorems, and are thus thought to represent the truth of some matter. The pragmatic theory of truth is also partly right, insofar as testing a proposition to see whether it is true or not is a practical matter of consequences of actions. And yet in one way the traditional pragmatic theory is misguided, because it confuses standards appropriate for a critical discussion or inquiry with those of a deliberation. Arguing that a proposition is true because it has good consequences, (or false because it has bad consequences), commits the fallacy of *argumentum ad consequentiam*.

The SE&D model can help us to understand concern for truth as a value. Concern for truth requires following the conversational postulates of being relevant, clear, and so forth. Persuasion dialog has rules, and collaboratively following those rules is evidence of a concern for truth. When both participants respect these rules, and use strong arguments and critically probing questions to interact argumentatively with the opposed side in a persuasion dialog, the outcome is not possession of truth but one of increase of the maeutic function. By being shown the weaknesses in one's own arguments, perhaps even being shown that they are fallacious, one comes to a deeper understanding of the evidential reasons behind a proposition at issue. By this means, both participants can gain a deeper understanding of the issue itself. This revealing of implicit assumptions is a kind of learning or increase of understanding. As a side effect it also produces toleration of opposed arguments that are shown to be not as unreasonable as they may have seemed at first. The outcome is a movement away from a rigid, absolutist view of an absolutistic kind that never admits defeat of one's own argument toward a more flexible view consistent with the proper use of defeasible argumentation.

Thus even though, following the view of Carneades, it might be unrealistic to expect that one would possess the truth, for now and for all time, at the conclusion of a persuasion dialog, it is still possible to see such a dialog as aiming at the truth

if the participants as rational agents have a respect for the truth that guides their argumentation. This doesn't mean that the participants always have to tell the truth, for based on all the considerations we have advanced so far, there is no reason at all to think that either one of them knows the truth, at least beyond all doubt. Indeed, the purpose of the persuasion dialog, or any type of argumentative dialog, is to confront doubts that exist or have been expressed, and address them not by removing all possibility of doubt, but by comparatively evaluating the arguments on both sides of the issue. Respect for the truth means acknowledging defeat if the other party has presented arguments that show that one's own argument is no longer tenable. In a persuasion dialog, respect for the truth needs to be defined in terms of defeasibility of argumentation and in terms of the rules and argument structures appropriate for a dialog.

10   Conclusions

It has been claimed that argumentation theory is a discipline that verges on a worrisome kind of relativism that tailors argument evaluation to audience acceptance. Boger (2005) put forward a sustained and well-supported argument alleging that, by subordinating truth to acceptance in this way, informal logic (argumentation theory) is infected with the "virus" of a pernicious relativism. This paper refutes the charge, by showing how a persuasion dialog can lead toward the truth of the matter being discussed, even if it does not resolve the conflict of opinions by proving beyond doubt that the proposition advocated by the one side is true and that of the other side is false. The evaluation of how a critical discussion leads toward the truth is evaluated by a determination of the depth of its argumentation, judged by six factors. We can still have a rational discussion of value issues, even if we are not in a position to end it by proving conclusively that our claim is true or that the opposed claim is false. Such a persuasion dialog need not subordinate truth to acceptability.

Definitions of the key terms persuasion dialog, critical discussion, defeasible reasoning, and depth of a discussion were given. Based on these definitions, and the argumentation technology provided by VAF's and the Carneades system, ten main conclusions about argumentation in persuasion dialog were drawn.

(1)   A persuasion dialog, as well as other types of dialog, like a scientific investigation, moves through stages, aiming toward finding out about the truth of a matter being discussed or under investigation.

(2)   Persuasion dialog can be successful even if the originating conflict of opinions is not resolved, provided there is a maieutic outcome characterized by an in-

crease in the depth of argumentation, revealing the strengths and weaknesses of the arguments on both sides.

(3)   The way a dialog or investigation moves forward is through a process of generating hypotheses that are tested and refined through interactively engaging with opposed argumentation and criticisms, producing stronger and deeper arguments.

(4)   Achieving depth and meeting standards of successful persuasion are different matters from determination of which side won the discussion by persuading the audience that its argumentation is more rhetorically persuasive than that of the opposed side.

(5)   At the closing stage of a persuasion dialog, the argumentation supporting or refuting the proposition accepted or rejected is no longer defeasible in that dialog, but it is not regarded as proving the proposition beyond all further doubt.

(6)   Still, the dialog can be closed off, and the ultimate conclusion it has led to can be accepted, on a balance of considerations, if it meets the appropriate proof standard.

(7)   Although we as rational agents in an investigation or discussion cannot know truth, or be sure beyond doubt we have it, still, using the evidence we have collected, we can still reasonably accept or reject propositions on an issue based on reasons.

(8)   Moving toward the truth in a persuasion dialog is evaluated through a process of collecting and testing of evidence, leading to rational acceptance or rejection.

(9)   On the SE&D model, the extent to which a persuasion dialog can be judged to have moved toward the truth can be judged by evaluating the depth of the discussion and the standard of evidence required to prove a claim.

(10)   Truth is an ideal, and respect for truth by both parties in a dialog is a value that can support the goal of resolving a conflict of opinions by rational argumentation.

Some of these ten theses are already widely accepted in working argumentation systems used in artificial intelligence, while others are more controversial. Together they present a philosophical viewpoint on argumentation that can be summed up as follows.

On this viewpoint, it can be conceded that it might be fair enough to say that there is a kind of "relativism" about what argumentation can achieve in a persuasion dialog. However, it is not a kind of relativism that holds that an argument is rationally correct merely because it persuades a particular audience to accept it. Instead, arguments are evaluated as stronger or weaker as they are processed through a sys-

tem like Carneades that judges the tenability of defeasible argumentation by standards appropriate for a dialog. It was shown that defeasibility is characteristic of the kind of reasoning used in persuasion dialog, and that arguments can be tentatively accepted, based on evidence, but should be subject to potential defeat in the future as new evidence comes in. On this Carneades-style view, truth and acceptance need to be combined in persuasion dialog. It is not committed to a kind of relativism that is pernicious because it subordinates truth to acceptance. It is based on the current understanding of how defeasible reasoning works as evidence in artificial intelligence, necessary for any systematic attempt to analyze and evaluate argumentation in real cases of doubt and disagreement. It steers a middle course between dogmatism and skepticism. It is based on a moderately skeptical view about our ability to possess the truth on matters of values, and is opposed to the dogmatic view that any agent using rational argumentation can claim an exclusive right to know the truth beyond all doubt. For part of the willingness to open a persuasion dialog and engage in rational argumentation in it should include admitting defeat when one's own argument has been shown to fail when confronted with criticisms or counter-arguments that present evidence showing it is no longer tenable.

Mutlti-agent dialog systems

In chapter 4, an outline of recent developments on communication techniques in multi-agent systems is presented. The developments are changing rapidly, and the outline given here only attempts to give the reader a feel of what the basic concepts and tools are, what has been achieved, and the direction new research is taking. It will quickly become apparent that the central problems that the developers of communication languages for multi-agent systems are dealing with are very similar to the dialog rules and principles discussed by Hamblin thirty years earlier (as outlined in chapter 2). Hamblin proposed a theory of formal dialogs in terms of what he called commitment-based formal dialectic. Some thirty years later, his commitment-based approach is becoming more and more favored by many of those working in the field of multi-agent systems. It is up to them to set the standards for agent communication, and to implement the technical requirements for the asking of questions, the framing of replies, and the other kinds of speech acts important for agent communication. These standards are based on a delicate balance between the communicative needs of the agent and the ability of the receivers to compute the meaning of the message within a "tractable" time (Dignum and Greaves, 2000, p. 1). The problem is to integrate the most practical and efficient ways to structure such communications technically, but that is also a problem of dialog theory.

In addition to simply presenting a survey of the recent work on multi-agent communication in computing, an attempt is made to show how this work fits into the field of dialog theory. It is argued that dialog theory is a useful fit as an overarching theory for this field that makes the various pieces of research into a unified structure. By offering a general rationale for how different kinds of argumentation should be structured in different kinds of dialog, insight is given into how rules of agent communication should be configured. Rules of agent communication that formerly seemed arbitrary or even confusing can begin to make sense as having a distinctive function when used for a communicative purpose. Thus dialog theory offers a normative structure in which problematic principles and rules of agent communication can at least be put into a logical and philosophical perspective that

brings them together into a unified framework as representing communicative actions that are used in different types of dialog.

1   Agent communication systems

In the past few years, a field of distributed computing called multi-agent systems (MAS) has grown rapidly (Wooldridge, 2002). Its applications include electronic commerce, manufacturing, and cooperative engineering. The concept of an agent has already been introduced in chapter 1, but it is worth quoting a couple of definitions of the term 'agent' by leading researchers in the field of multi-agent systems. An intelligent agent is defined as an entity (typically a software entity) that can carry out goal-directed actions in an autonomous way, according to (Wooldridge and Jennings, 1995). Another aspect of an agent emphasized in the definition given by Jennings, Sycara and Wooldridge (1998, p. 2) is that of "situatedness". Having his property means that an agent receives sensory input from its environment, and can also change this environment through its actions. Based on this notion, the definition of 'agent' given by Jennings, Sycara and Wooldridge (1998, p. 2) reads as follows "an agent is a computer system, situated in some environment, that is capable of flexible autonomous action in order to meet its design objectives". Ferber (1999, p. 9) proposed a ten-point minimal definition that can be summarized as follows.

1.   An agent is physical or virtual entity.
2.   It is capable of acting in an environment.
3.   It can communicate with other agents.
4.   It is driven by goals or objectives.
5.   It possesses its own resources.
6.   It can perceive its environment.
7.   It has only a partial and incomplete perception of its environment.
8.   It has skills and can offer services.
9.   It may be able to reproduce itself.
10. Its actions tend towards fulfilling its goals, depending on its perceptions.

Defining exactly what an agent is, in a fixed way, is not really possible, because the field of agent technology is in a rapid period of growth and development. But the definitions offered by the leading experts give insight into what the term should generally be taken to mean. A nice definition given by Singh (1998, p. 40) is: "In the true sense of the word, an agent is a persistent computation that can perceive its environment and reason and act both alone and with other agents." It might appear a bit strange at first to say that an agent can reason, because it is (typically) a software entity. Also, because it is programmed, it seems strange to think it is

somehow autonomous. The notion of autonomy is hard to define as well. What it means is that the system should be able to carry out actions without the intervention of its designers, so that it can be thought to have some control over its actions and even its goals (Jennings, Sycara and Wooldridge, 1998, p. 8). Yet another property of an agent is that it can perceive and understand the actions of other agents, and even collaborate with other agents, for example, in a deliberation dialog where planning requires teamwork.

Practical reasoning is the inferential process of arriving at a conclusion to take action based on an agent's knowledge of its perceived circumstances (Bratman, 1987; Searle, 2001). Woodridge (2002, p. 66) classified practical reasoning into two components: (1) the process of deciding what state of affairs to achieve, called deliberation, and (2) the process of deciding how to achieve these states of affairs, called means-end reasoning. The outcome of means-end reasoning is a plan (Bratman, Israel and Pollack, 1988), and thus practical reasoning is closely related to planning of the kind discussed in chapter 1, a field well developed in AI (Wooldridge, 2002, p. 66).

The basic argumentation scheme represents instrumental practical reasoning of a kind that is narrower than a derived kind of practical reasoning based on values (Clarke, 1985; Bratman, 1987; Audi, 1989). In the pair of schemes below, 'I' represents an agent and A represents an outcome of an agent's action.

> Basic scheme for practical reasoning
>
> I have a goal G.
>
> Bringing about A is necessary (or sufficient) for me to bring about G.
>
> Therefore, I should (practically ought to) bring about A.
>
> This form of reasoning can be undercut or defeated by asking any one of the following critical questions in a dialog.
>
> Critical Questions for Basic Scheme for Practical Reasoning
>
> (CQ1) What other goals do I have that should be considered that might conflict with G?
>
> (CQ2) What alternative actions to my bringing about A that would also bring about G should be considered?
>
> (CQ3) Among bringing about A and these alternative actions, which is arguably the most efficient?
>
> (CQ4) What grounds are there for arguing that it is practically possible for me to bring about A?
>
> (CQ5) What consequences of my bringing about A should also be taken into account?

In addition to the instrumental basic scheme for practical reasoning, there is a value-based scheme that takes into account values of an agent that may need to be considered.

Scheme for Value-based Practical Reasoning

I have a goal G.

G is supported by my set of values, V.

Bringing about A is necessary (or sufficient) for me to bring about G.

Therefore, I should (practically ought to) bring about A.

This scheme is based on a distinction between goals and values (Atkinson, Bench-Capon and McBurney, 2006, p. 164). In their formal analysis of practical reasoning, values are not just special kinds of goals. Values are different from goals in that they provide the reasons that an agent has for wanting to achieve a goal.

Critical questions for value-based practical reasoning

(CQ1) What other goals do I have that might conflict with G?

(CQ2) How well is G supported by (or at least consistent with) my values V?

(CQ3) What alternative actions to my bringing about A that would also bring about G should be considered?

(CQ4) Among bringing about A and these alternative actions, which is arguably the best of the whole set, in light of considerations of efficiency in bringing about G?

(CQ5) Among bringing about A and these alternative actions, which is arguably the best of the whole set, in light of my values V?

(CQ6) What grounds are there for arguing that it is practically possible for me to bring about A?

(CQ7) What consequences of my bringing about A that might have even greater negative value than the positive value of G should be taken into account?

For many cases, the basic scheme can be used to evaluate practical reasoning, but in cases of ethical, political or legal argumentation, like those typical in public affairs, values are important factors that need to be taken into account (Bench-Capon, 2002, 2003). To properly evaluate such cases, the scheme for value-based practical reasoning needs to be applied. Practical reasoning is a fundamental argumentation structure for multi-agent systems, especially in systems of electronic democracy, where the purpose of the system is to look at ways and means of carrying out political goals using intelligent deliberation in a democratic system (Girle, Hitchcock, McBurney and Verheij, 2003; Atkinson, Bench-Capon and McBurney, 2004).

For example, Gordon and Karacapilidis (1997) designed a mediation system to enable interested citizens and representatives of public interest groups to take part in electronic discussions with government officials planning public projects like setting zoning ordinances. The problem they encountered is that arguments expressed in it had to be represented in a simple enough way that laypersons could use it. Atkinson, Bench-Capon and McBurney (2004) have devised a structure called the Parmenides System that can be used by a democratic government to solicit arguments on a particular policy being contemplated. The system solicits criticisms of the policy by allowing justifications of a proposed policy to be presented, and allows critics to present objections to actions and goals. This capability brings us to the subject of agent communication.

In order to reason and act with other agents (as shown in chapter 1), agents have to communicate. At a technical level, agents transport messages over a network using a lower-level protocol like HTTP (Labrou, Finin and Peng, 1999, p. 46). But agents do not just send single messages to each other. Their communications take a conversational form in which each agent takes a turn sending a message related to a prior or next message by the other agent. Agents need to engage in dialogs to question other agents, to inform other agents, to negotiate with other agents, to monitor other agents, and to find agents that can help them to carry out tasks. Establishing a standardized agent communication language (ACL) is therefore a highly important part of MAS development. The first significant widely established attempt was KQML (Knowledge Query and Manipulation Language) proposed by the DARPA (Defense Advanced Research Projects Agency) knowledge-sharing effort. The central concept of this effort was that knowledge-sharing requires communication (Labrou, Finin and Peng, 1999, p. 46). KQML messages were defined by parameters that had various meanings. For example the content parameter referred to the content of the message, while the sender parameter referred to the sender of the message, while the receiver parameter referred to the intended recipient of the message (Wooldridge, 2002, p. 171). Several different versions of KQML were proposed in the 1990's. Examples of different kinds of KQML messages are given by Wooldridge (2002, pp. 172–175)

The most recent effort is the FIPA (Foundation for Intelligent Physical Agents) initiative (http://www.fipa.org/). FIPA is a non-profit organization with the goal of promoting the success of agent-based systems and their applications by making specifications and setting standards that will maximize the interoperability of such systems. FIPA operates through the open collaboration of its many member organizations all over the world. It will be through FIPA that standardized agent communication protocols and message formats will come into use in multi-agent systems. The standardized ACL developed by FIPA is similar to KQML. For example, in an 'inform' message, the sender might be identified as agent1 while the re-

ceiver might be identified as agent2. The content of the message might give the price of a particular commodity the two agents are discussing. This kind of message might occur, for example, in a negotiation dialog taking place in an electronic auction. In this example we see that the purpose of the message is for one agent to send information to the other, and therefore one might think that the dialog is a species of information-seeking dialog. But notice that more generally speaking, the transaction is part of the negotiation dialog. Later, in chapter 6, we will study how one type of dialog can be embedded in another. The FIPA communication language is built on what are called performatives. An example would be the accept-proposal performative, which allows an agent to state that it accepts a proposal made by another agent. Wooldridge (2002, pp. 176–178) has presented a list of these performatives. Another interesting example is the not-understood performative, which one agent uses to tell another that he has not understood some previous message.

As shown below, the current MAS initiatives to establish a standardized ACL have centered around information-retrieval systems in which an agent or group of agents is instructed by a human to collect information. Understandably, the focus of much of the attention so far has been on the giving of information by one agent to another, or on the asking of a question to get information. This point of view emphasizes information-seeking dialog, perhaps at the expense of the other types of dialog like deliberation and negotiation, now recognized as vitally important. But as will be noted later, there is much other research in computing that is also leading towards a recognition of the importance of deliberation and negotiation, as well as persuasion dialog.

ACL's are all dependent on a general theory of agency that lends some clear structure to defining not only what an agent is, but also what an action is taken to be. Defining what an action is in clear terms might seem trivial. But anyone who is familiar either with the recent computational theories of agency (Ferber, 1999, pp. 166–222) or with the earlier philosophical literature on action theory (Goldman, 1970) will know that it is not trivial and can be quite tricky. A theory of agency is "a general formal model that specifies what actions an agent can or should perform in various situations," including an agent's reasoning strategy (Dignum and Greaves, 2000, p. 4). Theories of agency used in ACL's are usually based on a small set of primitives characterizing certain mentalistic properties, especially belief, desire, knowledge and intention of an agent. Such mentalistic models are called belief-desire-intention or BDI models. The FIPA language is based on a set of BDI primitives, while KQML is based on a simpler theory of agency in which assertions are added to a knowledge base or deleted from it. As shown below, the latter approach is very similar to the commitment model.

2   Speech acts

Agent communication languages, including KQML and those supported by FIPA, are based on speech acts. Speech acts are also often called performatives (as illustrated in the previous section) or communicative acts in the ACL literature. Such a speech act occurs whenever one agent sends a message to another agent. The literature covers various speech acts that are perceived as being useful for agent communication. The main various kinds of speech acts that are important for most computing uses have been summarized nicely in a list of six categories with examples by Singh (1999, p. 8). The function of each speech act is indicated in parentheses, as indicated in the comparable list of types of performatives in (Singh, 1998, p. 43). For example, the function of the speech act of asserting (assertive type) is to inform the hearer about the statement asserted. Thus if I assert the statement 'The door is shut.' The function of this speech act is to inform you (the hearer) that the door is shut.

> *Assertive* (inform): The door is shut
> *Directive* (request): Shut the door – or query: Can pelicans fly?
> *Commissive* (promise): I will shut the door.
> *Permissive* (give permission for an act): You may shut the door.
> *Prohibitive* (ban some act): You may not shut the door.
> *Declarative* (cause events in themselves): I name this door the Golden Gate.

It may be interesting to note that in the earlier list, Singh also added a category of speech act called *expressive*. The function of the expressive speech act (Singh, 1998, p. 43) is to "express emotions and evaluations", as in the example, "I wish this door were the Golden Gate." This expressive speech act may not be as important in computing as the other six kinds of speech acts, but it could be important in representing the emotive uses in loaded questions of the kind connected with question-asking fallacies in chapter 4.

Some ACL's do not have the full range of speech acts of all the kinds indicated above. KQML has only speech acts that take the form of assertives and directives (Finin, Fritzon, McKay and McEntire, 1994). Agents can use these speech acts to tell facts, ask questions, subscribe to services, or contact other agents. The syntax has a simple form called a balanced parenthesis list. The first element in the list is a performative and the remaining elements are the performative's arguments (Finin, Fritzon, McKay and McEntire, 1994, p. 459). The so-called "ask-tell" message is often cited as characteristic of agent communication in KQML. For example one agent might send the message (ask, price IBM) and the other may respond with some figure representing the current price of an IBM share. The basic communicative structure presupposed is that you have two agents, and each agent has a knowl-

edge base. Asking a question represents the one agent's attempt to extract some item from the other agent's knowledge base. Thus the communicative context presupposed could perhaps be described as a kind of information-seeking dialog. In general, the development of speech acts in multi-agent systems has concentrated on this apparently simple type of information-seeking question. The two most important communication primitives in FIPA languages are INFORM and REQUEST, and all the other speech acts in FIPA can be defined in terms of these two (Wooldridge, 2002, p. 179).

One might expect more variety in the kinds of speech acts studied when more types of dialog come to be seen as important, aside from information-seeking. Pitt and Mamdani (1999, p. 335) expressed agreement with this approach when they made a comment specifically about FIPA, "The protocols specified by FIPA ACL can give purpose and meaning to individual performatives but only in the context of a conversation that has an identifiable objective". Adopting the perspective of dialog theory, the same performative can have a different function depending on how it is used in a different communication and what goal of dialog it is supposed to contribute to in that communication. The program of dialog theory is to study these different functions in different types of dialog.

To follow out this program of placing speech acts within the framework of dialog theory, it is necessary to adopt a thesis concerning speech acts and moves in a dialog. They need to be seen as equivalent or matching. In other words the thesis proposed here is that speech acts should correspond to the types of moves in a dialog. The speech act represents a kind of move that is made at a local level in a dialog. For, example, consider the speech act of assertion. It sounds like a very simple speech act. In MAS literature, the speech act of asserting is taken as an act whereby one agent informs another agent that a statement is true. As Singh defined assertion as a speech act above, to assert is simply to inform. Thus the speech act of assertion is simply a communicative act in which the proponent (speaker) makes a statement, representing an item in her knowledge base, and then passes it on to the respondent (hearer). Presumably, the statement was not already in the respondent's knowledge base. And so, it is inserted in it when the speech act of assertion is carried out. And that is that, or so it seems. So conceived, the speech act of assertion appears to be perfectly clear, and can easily and straightforwardly be defined as an act of transfer of information.

However we already know from chapter 1 that this apparent simplicity and clarity of the speech act of assertion is illusory. All is not what it seems. Yet, in the simplest kind of information-seeking dialog, assertion can be so simply defined. But communication of the information-seeking kind, particularly in its more refined forms, like expert opinion dialog, has a peirastic component. Once the peirastic component is introduced, then persuasion dialog also has to be considered.

And here, assertion is not as simple as it seems, when it comes to thinking out what kinds of rules and principles should be applied. In practice, assertion is connected to persuasion dialog, even in cases where the dialog is primarily one of information-seeking. Suppose, for example, that the case is one where an expert is being interviewed in order to give scientific information to a layperson (in the domain of expertise) in order for the layperson to take action. The expert could be a financial consultant who knows about investments, and the layperson could be trying to decide how to invest his retirement funds. In such a case, the basic framework is that of an information-seeking dialog. The questioner is trying to get information out of the expert that is relevant to his problem. The kind of argumentation used in such a case takes the form of the appeal to expert opinion. But now, look at the reality of what has to take place for the dialog to be successful. It is not just a simple case of the expert passing along statements, one by one, in response to the layperson's requests for information by asking simple factual questions. The expert's assertions might tend to be quite complex, full of qualifications and language that is hard for the layperson to understand. The layperson needs to not only ask for explanations and clarifications, but even to probe critically into what the expert says by asking appropriate critical questions. Thus to make the information-seeking part of the dialog successful, it has to be at least partly a kind of examination or peirastic dialog. In other words, adjoined to the information-seeking, there has to be another type of dialog with a critically probing component. This other type of dialog would seem to be a persuasion or critical discussion type. The point to be made here is that speech acts of asserting and questioning may not be as simple as they appear, in any kind of realistic case of information-seeking communication. In the background, persuasion dialog or other types of dialog may be involved.

It can even be seen that the speech acts of asserting and questioning are closely related in persuasion dialog. In the general type of persuasion dialog called permissive persuasion dialog in (Walton and Krabbe, 1995), there is a rule regarding assertions that is worth noting in this connection. Rule 9 for PPD (p. 137) is expressed as follows.

> *Rule 9*: If a party makes an assertion, the other party, in the very next move, is to make clear its position with respect to that assertion.

What Rule 9 means, in effect, is that the making of an assertion in PPD has a function comparable to that of the speech act of asking a question. When the proponent makes an assertion by stating a proposition *A*, according to Rule 9, the respondent must at the very next move reply by saying that he either accepts *A* or not. This is a fairly aggressive rule regarding the speech act of assertion, and how that speech act is implemented in a dialog. When the proponent makes any asser-

tion at all, the respondent cannot just ignore it and move on. He has to declare his agreement or disagreement right away. Is this rule reasonable? The answer is that it seems a little aggressive, as a rule to have for persuasion dialog generally. But it does perform a function in PPD by making commitments come out explicitly as the dialog proceeds. This function is quite valuable in PPD, because the central problem in this type of dialog is to get the respondent to declare explicit commitments, instead of always being able to evasively reply, "No commitment." So even from this brief consideration of Rule 9, it becomes apparent that assertion, and the rules and principles surrounding it, are linked to the problem of retraction. An even bigger question is whether assertions can be retracted.

Once a party has asserted a statement, and thereby become committed to it, can she then retract it? This issue is really the central question in structuring persuasion dialog as a whole. The issue is treated extensively in (Walton and Krabbe, 1995), and little more needs to be said about it here. In general, the principle is that a party should be free to make retractions, but not universally free. If she retracts one statement, she may have to retract others as well, or she may be open to penalties, or attacks of various kinds (Walton and Krabbe, 1995). Although persuasion dialog must, by its nature, be fairly permissive in allowing for retractions of commitments, other types of dialog are not so permissive. For example, the inquiry type of dialog is cumulative, meaning that once a statement has been accepted, it should never (ideally) be retracted. Deliberation dialog, like the persuasion type of dialog, needs to allow for considerable freedom for retraction of commitments. The reason has to do with the nature of deliberation. It is typically carried out under conditions of uncertainty where new information is rapidly coming in. Deliberation needs to be flexible, to adjust to rapidly changing conditions. The inquiry, on the other hand, needs to be stable. Continual retraction of previously accepted statements would be disruptive, and would go against the whole aim of the process, which is to build on solid foundations.

3   Interrogative messages in ACL's

The speech acts are implemented in ACL's by programming them in terms of preconditions and postconditions that are expressed using variables. A simplified version of an example from (Dignum and Greaves, 2000, p. 6) shows how this feature works. For example, suppose the receiver has a set of numbers as its knowledge base and the sender asks some question about the numbers. A precondition might be that the question is expressed in such a way that numbers can be used to answer it. A postcondition might be that the question is answered successfully if a number is given as a reply to the question. Thus the speech act is represented in an ACL so

that it can be programmed using preconditions and postconditions. For programming purposes the conditions need to be formulated in a way that is precise, and that is not too costly to implement in computer time. Thus naturally a simple formalistic version of the speech act that is not too cluttered with complexities tends to be more useful for a programming purpose. The preconditions and postconditions are expressed formally by variables and their values before and after the action takes place. But at a higher level of abstraction, they are formulated through the theory of action adopted (Dignum and Greaves, 200, p. 6). Thus in a BDI theory, the preconditions and poscondtions will be xpressed in terms of the beliefs, desires and intentions of the sender and receiver.

In KQML, there are three kinds of interrogative messages that can be sent by one agent to another: ask-if messages, ask-one messages and ask-all messages. One agent will send another an ask-if message if the first agent wants to ascertain whether a specific item of information is in the knowledge base of the second agent. For example, the one agent can send the message 'ask-if (Tucson is south of Phoenix)' to the other agent. In reply, the receiver agent should tell whether the statement 'Tucson is south of Phoenix' is in its knowledge base or not. There are only two possible replies to an ask-if question – in effect, either a 'yes' or 'no'. Thus an ask-if message is equivalent to what is often called a yes-no question. One agent will send another agent an ask-one message to get a single match, and an ask-all message to get all matches in the knowledge base. For example, an agent could send the message 'ask-one (city with over ten million population) or 'ask-all (city with over ten million population)'. An example of an ask-one question is a query about the price of IBM stock from agent Joe (Labrou, Finin and Peng, 1999, p. 47). The content of the message would be: (PRICE IBM?price), and the so-called ontology would be NYSE-TICKS. The ontology indicates that the recipient of the message is a stock-server. The content indicates what information is required to satisfy the question. In this case, it is a current price for IBM stock.

Two speech acts that have been prominent in structuring interrogative messages in ACL, as noted above in connection with FIPA, are 'request' and 'inform'. Cohen and Levesque (1995) use these speech acts to specify the form of an ask-if message (yes-no question). Where A is a statement variable and K is the message recipient's knowledge base, this type of message could be said to have the following form: request ((inform S is in K) or (inform S is not in K)). Thus it would seem that the set of speech acts, along with some logical operators like 'and' and 'or', could be used to define and classify the various types of interrogative messages used in ACL's. This analysis is very promising, in light of the discussion of questioning in chapter 4, because it ties the definition of the type of question to how such a question would or should be used (the function of the question) in a type of dialog. The yes-no question, according to this account, has the function of a request for information,

presumably within the context of some sort of information-seeking dialog. This account does seem to define the basic function of a yes-no question in the simplest and most basic type of information-seeking in just the right way. But as indicated in chapter 4, matters are rarely this simple in real argumentation.

The FIPA ACL provides an additional dimension to interrogative message structure. Suppose the recipient agent does not understand the message. For example, the message may be in a format that the recipient does not read. Or it may be a message type that the recipient agent can read but cannot respond to. The recipient has the option of not accepting the message. For example, it can respond with a 'not-understood' message. The recipient agent can also add a reason why the message is not understood. Thus the FIPA ACL has some capacity to deal with communication failure and error management in handling interrogative messages. This additional feature could be quite useful, because it could leave the respondent a way out, not only in dealing with questions that are not in the right syntactic form, but even in dealing with argumentative and loaded questions of the various problematic kinds considered in chapter 4. But still, it does not really seem to be designed to deal with this order of sophistication in questioning practices.

Game theory is an approach that has often been used in multi-agent systems as a way of reaching agreement in negotiations and other types of dialog, but it has a number of disadvantages (Wooldridge, 2002, pp. 148–149). One is that it cannot capture the notion of a position being justified by reasons offered to support it. Another is that positions cannot be changed, whereas as we have seen in chapter 1, retraction of commitments is possible in dialog systems. These limitations of game theory have led to new multi-agent systems based on argumentation, as shown in (Wooldridge 2002, pp. 154–155). Multi-agent systems have now recognized types of dialog other than negotiation, information-seeking and deliberation, including persuasion dialog and inquiry.

Existing practices of the kind considered in chapter 1 make it clear that the kind of argumentation characteristic of practical communication is conversation-oriented. The way agents communicate in practice is implicitly structured around certain types of conversational interactions. However, the properties of these organized types of conversations between agents, and how the speech acts should best be supposed to fit into them, has not yet been systematically enough investigated in multi-agent research. Some steps have been taken, but what needs to be done is to see each speech act as having a different function, depending on what type of dialog it was used in. For example, consider the speech act of assertion. As indicated above, Singh defines the speech act of asserting (assertive) as having the function of informing. This account of the speech act of assertion is right for the simplest kinds of cases, but is narrow when more complex kinds of communication, like expert opinion dialog, come to be considered. As shown above, asserting has a different function in persua-

sion dialog, where it is linked to the global goal of the critical discussion as a type of dialog. Of course, in information-seeking dialog, assertion has a similar function, and it also involves making a commitment to the truth of the proposition asserted. But the way this commitment is made is different in each different type of dialog. In both types of dialog, an assertion can be retracted. But the conditions for retraction will be different, because the goal of the dialog, and the arguments used in that dialog, are different. Thus assertion has a common structure as a speech act in each type of dialog. But exactly how that speech act will be used, and its function, will vary from one type of dialog to another.

Research on ACL's has already gone well beyond consideration of single speech acts and is beginning to study how each speech act functions in each different type of dialog. Dialog theory offers a philosophical perspective to give this research a new level of generality and direction. The usefulness of dialog theory is that it can link up the extensive work already done on speech acts and carry it to a new level that is more general, and at the same time much more applicable to the needs of ACL's. Considerable work that leads in this direction has already been done, indicating the importance of this direction for ACL's. So-called conversational policies have been studied to try to define the conditions under which communicative acts can properly be used.

4   Conversation policies

In order to try to make the speech acts in ACL's work more efficiently for communicative purposes, more general communication principles called conversation policies have been added. An example from Arcol, cited above, is the policy that a sending agent can only tell a receiving agent something if the sender already believes it and the receiver does not believe it. Another example is the INFORM conversation policy (Smith, Cohen, Bradshaw, Greaves and Holmback, 1998). According to the INFORM policy, when a sender informs a receiver that some statement A is true, the receiver either acknowledges or remains silent. Expressing this policy in a BDI framework, INFORM can be read as saying "I, the sender believe that A is true." ACKNOWLEDGE can be read as saying, "I, the receiver believe that you the sender believe that A is true." The conversation policies in FIPA are based on a BDI framework in which both sending agent and receiving agent have so-called "attitudes" of belief, knowledge and uncertainty. In this framework, belief and uncertainty are mutually exclusive, while knowledge is defined as an abbreviation for belief or uncertainty (Pitt and Mamdani, 1999, p. 336). This way of defining knowledge is typical of how the BDI framework is implemented in knowledge-based computer systems, but it is highly atypical of how knowledge is viewed

in philosophical logic. In philosophical logic, to say that a proposition is known to be true is taken to logically imply that it is true. This sense of the term 'knowledge' is often called Platonic, following Plato's theory that knowledge is about the true and unchanging reality. If something is true, on this view, it could never turn out to be false later on. Thus underlying the formulation of conversation policies in multi-agent systems may be different interpretations of what knowledge and belief should really refer to as mental states or attitudes.

Expressed in BDI terms, the conversation policy for INFORM in FIPA is described by Pitt and Mamdani (1999, p. 336) as based on "feasibility preconditions" of the speech act of informing in FIPA. The first feasibility condition for the speech act of informing in FIPA, in their terms, states simply (as indicated above) that the sending agent believes that the proposition contained in the speech of informing is true. But there is also another more complex precondition. It can be explained by adapting the notation of Pitt and Mamdani slightly for purposes of exposition. Let us designate the proposition contained in a FIPA INFORM message by the letter A. The second more complex feasibility precondition of INFORM in FIPA can now be stated as follows (Pitt and Mamdani, 1999, p. 336): "The second feasibility precondition for INFORM states that the sending agent does not believe that the intended recipient either believes or is uncertain that either proposition A or its negation is true." The intent of this precondition is to eliminate the possibility that the sending agent might try to inform the receiving agent of some proposition A that he believes that the receiving agent already believes. As noted above, knowledge, according to the FIPA framework, is defined as an abbreviation for belief or uncertainty. Thus the second feasibility precondition could be equivalently expressed as follows: the sender does not already believe that the receiver has any knowledge of the truth of the proposition A. This formulation of the second feasibility precondition, expressed essentially in the form above by (Labrou, Finin and Peng, 1999, p. 49), makes more intuitive sense. It can also be seen to fall under the Gricean conversational principle to the effect that a collaborative participant in a conversation should not inform another participant of something the second participant already knows, or at least something the first participant thinks the second participant already knows. According to the maxim of quantity (Grice, 1975, p. 67), a participant is asked to "Make your contribution (to a conversation) as informative as is required." This conversational maxim would imply that a participant should not pass along information that she knows that the other participant already knows. Evidently the purpose of the FIPA precondition is similar to that of the Gricean conversational maxim. It is to disqualify as useless communication the kind of INFORM message in which the speaker sends the hearer information that is superfluous, or not needed, as far as the speaker knows. This precondition promotes more efficient communication by cutting down on the communicating of useless or irrelevant infor-

mation. From the viewpoint of dialog theory, it relates to the subject of relevance, especially of the kind important in information-seeking dialog.

What is of interest to note here is that these two feasibility preconditions seem to be related to principles that Hamblin discussed as possible commitment rules for formal dialectic. The first precondition states that the sending agent believes that the statement he informs about is true. This precondition looks close to some of the Hamblin principles discussed in chapter 4, but does not match any of them exactly. Hamblin's question rules do lay conditions on the commitments of the questioner. But his assertion rules do not state that the proponent must be committed to the statement she asserts. Why not? Presumably the reason is that this condition, in effect, states that the proponent of the assertion is not lying. Therefore this condition on assertion in formal dialectic would relate to the sincerity of the maker of an assertion. We return to the subject of sincerity below. Here we could note that for very simple purposes of knowledge-based systems in computing, it could be reasonable perhaps to just assume sincerity of a message sender. But for Hamblin's purpose of the study of argumentation generally, and especially for the study of fallacies, sincerity of the message sender could not be taken for granted.

The second precondition does seem to correspond fairly well to one of the principles that Hamblin discussed. The second INFORM precondition says that the sender of a statement does not believe that the receiver already knows the statement to be true. This precondition is reminiscent of Hamblin's Assertion Rule 1, which says that the proponent may not assert a statement if that statement is already a commitment of the respondent. Of course, the Hamblin rule is a commitment rule whereas the INFORM precondition was expressed (as above) in BDI terms. As noted in chapter 4, (Hamblin, 1970, p. 269), Hamblin even justified Assertion Rule 1 by noting that it seems reasonable if we regard the function of an assertion as the giving of information.

5   Sincerity conditions

The above two preconditions of the INFORM speech act are often interpreted in the literature on multi-agent systems as being based on, or as expressing presuppositions about the sincerity of the agents that are communicating with each other. One such pair of presuppositions that appears to be inherent in the INFORM policy are the assumptions that the sender is sincere and that receiver believes that the sender is sincere. In fact, the first feasibility condition above is often called the "sincerity condition". It would seem that, according to the sincerity condition the possibility of deceptive communication is simply being ruled out by assumption of the conversation policy that has been implicitly adopted. But studies on the falla-

cies in informal logic have shown that such a sincerity condition is a serious limitation on representing realistic communication attempts and strategies of argumentation. In *ad hominem* argumentation, and many other kinds of argumentation that depend on the credibility of the arguer, it would be far too limiting to have to always assume that the speaker is being sincere, in the senses represented by the feasibility preconditions above.

One might think that not always operating on the assumption that sincerity conditions hold may be characteristic of persuasion dialog, but that in information-seeking dialog, it could be generally reasonable to assume that the participants are sincere, and are not trying to deceive each other. Indications to the contrary have already been seen however, in chapter 4. It may be true that in the very simplest kind of information-seeking dialog, the sincerity of the sender or receiver of an INFORM message does not need to be considered. But as shown in chapter 4, once we take a realistic look at forms of argumentation like appeal to expert opinion, this assumption of sincerity must be regarded as open to questioning, as part of the dialog format of the argumentation. The question of whether a giver of information may be lying becomes an important part of the critical evaluation of the argumentation. For such purposes, sincerity of the proponent needs to be seen as an additional condition that can hold or not, in a given case.

This limitation shows how specialized the current ACL conversational policies are, because they are being applied to a specific use. It would be helpful if they could be rethought at a higher level of generality from a point of view of dialog theory. The sincerity of the speaker, for example, must be open to challenge in some cases of dialog, even though it can be assumed in others. As Pitt and Mandami (2000, p. 51) have noted, for example, many multi-agent systems are being developed for electronic commercial activity involving negotiation. In such a system, the notion of sincerity has to be limited. An agent does not have to tell the whole truth, and some degree of deception may be tolerated. Thus the conditions on an inform message should be different in a negotiation dialog than they would be in an information-seeking dialog. The limitations of seeing the various preconditions of the INFORM message in a BDI framework could also be overcome by reformulating in the clearer commitment framework of dialog theory. Viewed in light of dialog theory, the INFORM speech act should properly be taken to mean that the speaker is making a commitment to the hearer that the statement *A* passed to the hearer is contained in the commitment set or data base of the speaker. Of course, such a commitment could later be retracted should new information come in, as a dialog proceeds. Also, such an assumption of sincerity does not have to be universal. It could be adopted as a rule in some types of dialog, and not adopted as a rule in other types of dialog. Once we start looking at the preconditions as commitment rules of the Hamblin type, and stop having to see them in terms of complicated and mysterious iterations

of knowledge and belief, how to deal with them becomes a much more manageable problem. They can now be seen as commitment rules, and they can be adopted or not, depending on whether the function they perform contributes to a type of dialog or not. We can now tailor the specific preconditions to the requirements of the type of dialog, and the role performed by the precondition as a dialog rule.

Another condition on the speech act of informing is that the statement asserted by the proponent should be relevant information of the kind that the respondent requested or needs. At least, this is roughly what INFORM should mean, as expressed in a commitment framework, in an information-seeking dialog. In this framework, the rationale of the feasibility precondition begins to make sense, instead of being entangled in the confusing and apparently abstract and artificial BDI assumptions that are so hard to make sense of. Thus putting the various speech act conditions into the commitment-based framework of dialog theory does not solve all the problems posed by these conditions immediately. But by putting them into the dialog framework, it makes it possible to see how to go ahead to give a basis or rationale for accepting or rejecting a given condition as performing a function in a type of dialog. In other words, the problems are transposed into a different setting where what needs to be done is to look at the dialog structure as a whole, and see the place of the given condition in that overarching structure.

It should be noted here that there is a formal equivalence between BDI conditions and commitment-based conditions. So we are not trying to argue here that the analyses of the various speech acts using sets of BDI conditions are inherently wrong or incorrect. The question at issue is rather one of the naturalness of the theory that is employed to put the speech acts in a holistic philosophical perspective that makes sense intuitively and can be expressed formally in an over-arching structure. By this means one proposed set of preconditions and postconditions can be compared with another, and then a decision can be made on which is the more useful set to use for specific programming purposes in an ACL designed for some purpose. The argument here is that the commitment-based model is generally better because it is more comprehensive and more natural for most applications in ACL's, and because it has a dialog structure that is dialectical. There may be applications where the BDI architecture fits better and is more natural. But since the commitment-based model is more easily amenable to formalization, and is showing that it does not run into the intractable problems of obscure formulations faced by the BDI approach, it is preferable as a general theory. More and more researchers in multi-agent systems are leaning towards and advocating the commitment-based approach (Singh, 1998). A summary of the objections to using only the mental BDI concepts for specifying ACL semantics has been given by Singh (2000) in the following two basic points. First, communication is a public phenomenon whereas mental notions are private (p. 33). Second, under the cur-

rent theories of mental concepts, we cannot verify the beliefs and intentions of an agent even if we know how it was constructed (p. 34). These points support the argument that a commitment-based model represents a better direction for future work to take. In the end, perhaps, the BDI model may even be expressed within the commitment model in a satisfactory way. At any rate, the two models are closely related from a structural point of view, and it is not being argued here that the BDI model should be abandoned, or is worthless. Much the same phenomenon is happening in computing as in analytical philosophy in recent years. The analytical philosophers became dogmatically attached to the BDI model, even though it led them to intractable problems about justified true belief and other epistemological problems that could not be solved. Hence they kept circling around in the same swamp (as Michel Scriven once put it) without ever coming to some useful general theory of justification of belief and knowledge. Modal logics of knowledge and belief were helpful up to a point, but by themselves were just not rich enough, or even the right sort of structure to model speech acts and knowledge-based multi-agent communicative reasoning. But dialectic is the right sort of structure. Now that the computer scientists are starting to accept dialectic as a good direction to take, hopefully more of the analytical philosophers will follow.

6   Understanding of messages

Another important assumption of the INFORM conversation policy is that both the sender and receiver understand statement *A*. For example, the receiver must either acknowledge or remain silent. Acknowledgement appears to imply acceptance. The receiver does not have an option of acknowledging having received the message, but only of adding that the statement *A* is not expressed in a frame of reference that it can understand. Just as in chapter 1, it was shown that there are several rules in the critical discussion about agreeing to enter into the discussion at its various stages, the same principle must apply to information-seeking dialogs in ACL's. There should be a presumption that both speaker and hearer understand the statement that is passed along as information. But now some general questions loom large. What is "understanding"? What happens if one of the parties does not understand the statement? Should the other party, in such a case, be obliged to offer an explanation, or at least an explanation-attempt of the statement? Building in such an explanation feature seems like it could be extremely useful in ACL. In fact, the usefulness of the capability for explanations has already been proved in the communication languages developed for expert systems. But how could such a function be built into ACL's? One problem is that the philosophical work on explanation has been dominated by the older semantic model. What is needed is a new pragmatic model based on a no-

tion of understanding appropriate for the study of verbal communication (Dascal, 2003). Such a concept of understanding needs to be of a kind that could be used in different contexts of communication, like legal and scientific contexts, and that could be incorporated into the different types of dialog in different ways for ACL's so that a really useful INFORM policy could be worked out.

In the last half of the twentieth century, the dominant model of explanation was the covering law model, or deductive-nomological model, as it is also called. This model defines an explanation as a deductive (or inductive, in some instances) inference from a set of initial conditions and a set of general rules (universal generalizations) to a statement to be explained. However, recent work in AI on computational generation of explanatory dialog has often assumed a dialog model of explanation that is much richer than the covering law model. What would be helpful at this point is a logical framework in which the dialogical model can be expressed in a precise way so that it can be seen as a worthy competitor to the covering law model. To begin this task, a first attempt at a dialectical analysis of the concept of explanation has been put forward in two book chapters (Walton, 2004, chapter 2; Walton, 2005, chapter 6). The analysis consists primarily of a set of conditions defining what an explanation is as a type of speech act used in a dialog exchange between two parties. A more recent paper (Walton, 2007) takes on the task of providing a formal system of dialog in which the new dialectical speech of explanation is a main component. This system is an extension of CB called CE. In CE the speech acts of requesting an explanation and providing an explanation are represented as dialog moves that must conform to opening rules, locution rules, dialog rules, success rules and closing rules.

This dialectical theory of explanation postulates that requesting an explanation is part of a dialog between two parties in which one of whom asks a question concerning something which he claims not to understand. To provide a successful explanation, the other party must offer a response that provides the requested understanding to the party asking the question. The dialogical model of explanation is based on the assumption that the notion of understanding is clear enough to be a component in defining explanation. To help grasp this notion, we need to turn to recent work in cognitive science. Schank (1982) contrasted a strong kind of understanding called complete empathy, like that between twins, with a minimal kind of understanding he called "making sense". Making sense occurs where there is a gap in a situation that generally makes sense to us, but there is one particular point in which it fails to make sense – an anomaly or inconsistency. Dascal (2003, p. 304) also made the point that we have to approach the concept of understanding in a roundabout way by first of all trying to grasp the many ways misunderstanding can occur. Making sense is a repair process in which one party helps another to account for an apparent anomaly evident to the other by using scripts. Scripts represent knowledge

people are assumed to have about common situations, and knowledge they have about routine ways of doing things (Schank and Abelson, 1977). In the usual example, called the restaurant script (Schank, Kass and Riesbeck, 1994, p. 7), a person can be taken to know a set of routine actions and common expectations about what is or is not done when he or she goes to a restaurant, like paying for her meal.

An explanation is successful if it transfers understanding of what was asked about to the questioner who asked for an explanation, but how can we test whether understanding has been successfully transferred in a given case? Scriven (1972, p. 32) answered this question by suggesting that we test comprehension or understanding of something by asking the subject questions about it of a particular kind. These questions test not merely the recovery of information that has been explicitly presented. They must test the subject's capacity to answer new questions in a dialog. But what type of dialog contains such a kind of testing by questioning? The answer is that it is a species of information-seeking dialog called examination dialog. Examination dialogue is classified by Dunne, Doutre and Bench-Capon (2005) and Walton (2005, 152–153) as a species of information-seeking dialogue that sometimes shifts to a critical discussion in which the questioner finds inconsistencies and other inadequacies in the subject's collective replies (p. 1560). The questioner's goal is that of finding out a party's commitment on some subject being discussed and testing it against the known facts of the case. This kind of examination dialog is familiar to us in legal contexts in trials, and in the examination of students in educational contexts.

In CE, like CB, each participant in a dialog has a commitment set, and in a comparable way, a participant's understanding of what is being discussed will change and evolve over the course of the dialog. Some implications of this new pragmatic approach to the concept of explanation will be discussed in chapter 7. In recent ACL's some steps in this direction have already been taken, indicating the importance of the notion of the understanding of a message.

This limitation of having to assume that both speaker and hearer always understand a message has been overcome in the FIPA request protocol. This protocol allows an option whereby the receiver may respond to a message, like a request for information, by replying that it does not understand the request. According to the FIPA protocol, the receiver may have several choices of how to respond to a message. It can send "not-understood", "refuse", or "accept", upon initial receipt of a message. Also, the receiver of the message has three options – failure, inform-done, or inform-iota. "Failure" reports failure to act on the request. "Inform-done" reports action on the request. "Inform-iota" reports that the request has been passed on to another party called the "initiator". Failing to accept a message is a necessary option, but basing such a failure on a reason, like not understanding the message, could also be an extremely useful option to have in an ACL.

FIPA has also introduced an additional more complex message type called a "request for proposals". Included under this category is a special type of message called the "contract net proposal" which contains a temporal clause. In a contract net protocol, the so-called "initiator" sends out to multiple receivers a call for proposals (cfp) message for carrying out a task by a specified deadline. A receiver, called "the participant" can reply "refuse", "not-understood", or can offer a proposal. A proposal, for example, might be a price for carrying out the designated task by the stated deadline. The initiator can then accept or reject the proposal. It would seem that the request for proposals message is a different type of conversational action than the INFORM type of message. The underlying purpose of the message seems to be different. The request for proposals message is a request for the receiver to make a proposal to carry out some designated type of action. The purpose is to get action of some sort, to carry out a task or meet some kind of goal. The purpose of the INFORM message is quite different. It is to transmit information from a sender to a receiver. Of course, this message is a kind of action – a speech act. But its purpose is not to get action, as such. Its purpose is simply to communicate information by passing along a statement that the sender (supposedly) thinks is true. In short, one could distinguish quite generally between two different communicative purposes underlying conversational policies at a higher level of abstraction. This suggestion, like many others made above, indicates the need to move from single speech acts and localized messages to considering the types of dialog these messages and speech acts are part of.

7   Rational effects of a message

Above it was shown how the FIPA feasibility preconditions serve as conversation policies governing the conditions under which the speech act of INFORM can appropriately be sent by one agent to another. But it can now be added that FIPA has, as well as the preconditions, other conversation policies relating to the effects of sending such a message. These conditions relate to how a sending agent can anticipate what is called the "rational effect" of a speech act. According to Labrou, Finin and Peng (1999, p. 49), "A communicative act's rational effect represents the effect that an agent can expect to occur as a result of performing the action; it also typically specifies conditions that should hold true of the recipient". So defined, the rational effect is not something that must actually occur when a speech act is carried out. Two additional FIPA conversational policies define inferences an agent can make by observing the communicative effect of carrying out a speech act. These two FIPA policies are defined formally as properties by Pitt and Mamdani

(1999, p. 337), but the thrust of these formal definitions can be conveyed in the two conversational policies stated below.

> *Communicative Effect Policy 1*: if the sending agent sees that an action has taken place, and that the receiving agent carried out that action, then the sending agent can infer that the receiving agent had the intention to bring about the rational effect of the action.

> *Communicative Effect Policy 2*: if the sending agent sees that an action has taken place then that agent can infer that the feasibility preconditions of this action hold.

The structure of these conversational policies has to do with the sending agent's ability to draw inferences, particularly about the beliefs and intentions of the receiving agent, on the basis of observing the effects of a speech act. Such an inference involves feedback, and it involves the ability of one agent to judge the internal states of the other agent, based on these observations of what occurred in a dialog. Both policies seem to reflect the notion of agent communication portrayed by the simple example of plan recognition in chapter 1, section 3 (Huhns and Singh, 1998, p. 87). In this example, the one agent sees the other agent opening the refrigerator door, and infers that the other agent does not believe it is cold outside. In this example, the agent's inference is not drawn from a speech act, but simply from the observation of the other agent's action. Even so, the principle of how agent communication involves an inference drawn by feedback, from one agent's observations of the actions of another, is evident. These two policies bring out the lesson already formulated in chapter 1, section 3 that agent communication is harder to represent than you think as some sort of clear dialog. Once feedback becomes involved, and communication comes to depend on one agent's ability to probe into the beliefs and intentions of the other agent, the problem of how to formalize a conversational policy as a rule of dialog becomes acute.

Exactly what these two communication effect policies state, and how they should be implemented as rules of dialog, are matters that are not easy to clarify. Perhaps because they are expressed in mentalistic BDI terms, it seems hard to know how to implement them exactly as clear and precise dialog rules. One such problem has been explicitly formulated by Pitt and Mamdani (1999, p. 341). It could be called the color problem. Suppose that the sending agent believes that the color of some object is red. It informs the receiver that the color of the object is red. By communication effect policy 2, the receiving agent will infer that the sending agent believes that the color of the object is red. But then suppose that new information comes in, and the sending agent comes to believe that the color of the object is blue. The INFORM preconditions are satisfied, so the sending agent can

inform the receiving agent that the color of the object is blue. So according to com-munication effects policy 2, the receiving agent can now infer that the color of the object is both red and blue. The problem posed here is one of how to deal with new information that conflicts with previously held beliefs. This problem can be dealt with in various ways, but as Pitt and Mamdani note (p. 341), it is not covered by the FIPA semantics. The problem then is to devise some natural and illuminating way of dealing with dialog sequences in which new information leads to conflicts or other problems that call for the updating of beliefs.

Reed (2006, p. 24) explained that FIPA basically characterizes dialog in two ways. First, it specifies the conditions for a locution that must hold before it can be uttered, and for the update after the location is uttered. Second, it does this on the basis of the mental states of the interlocutors. For example, an inform utterance can be made only if the speaker believes the content. This way of proceeding limits FIPA not only to a BDI framework, but in a way that places emphasis on the design of clear and unambiguous communications of a simple kind. Basically, FIPA views messages as composed by modeling dialogs as finite state machines, in which each message is seen as a transition from one state to the next state. Because of this lim-ited approach, such models are not well suited to representing natural language con-versational dialog of a kind that one human uses in communicating with another, or even to representing communications between a human user and a computer.

8   Future multi-agent systems and dialog theory

As shown above, the current initiatives in ACL's have centered around considera-tion of conversation policies relating to the speech acts of questioning and assert-ing in the information-seeking type of dialog. What is shown is that there seem to be two different approaches to defining these speech acts. Each seems to represent a different philosophy on the way things should proceed. In a comment on this chapter (March 21, 2001), Chris Reed distinguished between two distinct ways of defining ACL primitives that have become apparent as the MAS literature has evolved. The one view takes the semantic definitions of the primitives as in part determined by the context of dialog. For example, on this view, 'inform' takes on one meaning in an information-seeking dialog, but has a different meaning in a negotiation dialog. On the other view, each primitive has a meaning that is inde-pendent of the context of dialog, and that is the same in all kinds of dialog. Part of Reed's commentary is quoted below, because it explains the difference between the two views in clear and precise terms.

It is emerging that there are two distinct views of ACL primitives within the ACL literature. By this view, the meaning of individual primitives –that is, the semantic definition of primitives by which they can be implemented – is at least in part determined by the context in which they occur. Thus an 'inform' will take on one meaning if it is used in a particular type of information-seeking dialog, and will take on a different meaning if used, say, during a negotiation. Looking at it the other way around, the definition of the primitive 'inform-in-information-seeking' involves, perhaps, sincerity on the part of the utterer and crucially, some component of what happens next – perhaps the hearer is obliged to acknowledge or indicate non-understanding. In contrast, the inform-in-negotiation might involve less than sincerity on the part of the speaker (say, nothing more than that the hearer end up believing that the speaker intended the hearer to believe the proposition in question) and, crucially again, require the hearer to, say, accept or reject the claim. The key point, by this first point of view, is that aspects of policy be specified as part of the operational semantics of primitives. The alternative point of view is to clearly separate out the two semantic components of agent language. By this second view, individual primitives have a meaning independent of the context in which they are used. This does not mean that there is necessarily only one 'inform'. Thus inform-in-information-seeking and inform-in-negotiation can still be distinguished. The difference lies in the semantic definition, which admits of no component of the protocol. So whilst it is acceptable to include the feasibility preconditions that involve, for example, reference to speaker beliefs and intentions, and rational effects that refer to hearer beliefs (and possibly, unusually, intentions too), it is inappropriate to include reference to what communicative acts should follow.

As Reed noted, KQML unwittingly subscribed to the first view. 'Unwittingly' is an appropriate term because, at that time, the distinction between the two views had not been drawn. He also noted that the second view integrates well with the commitment-based approach to dialog. This second view looks quite hopeful as fitting well with dialog theory, especially once ACL research has broadened out in the future, as it inevitably will, to take other speech acts and uses of questions in different kinds of dialog into account.

Taking an even wider look at AI research in recent years, the use of dialectical argumentation has had other directions and motivations that are worth noting. One has been the mass of work on negotiation dialog, stemming from many commercial applications. Another has been the use of argumentation in AI as a proof-mechanism for defeasible reasoning. These argumentation mechanisms are inherently dialectical, because the arguments against a claim need to be weighed against the arguments for it (Prakken, 1993). Application of these mechanisms to legal argumentation, for example, suggests that the persuasion dialog structure is important (Feteris, 1999). These various uses of the dialog theory approach to argu-

mentation in AI broaden out the perspective to include negotiation and persuasion dialog, as well as information-seeking dialog.

As Reed (2006, p. 24) pointed out, although there is a strong tradition of dialog design in multi-agent systems, the emphasis has been on simple means of characterizing communications between agents. On the other hand, the work in argumentation like (Walton and Krabbe, 1995), is at a level of abstraction that makes it underspecified, from a point of view of implementing it in a working computer system. So, much work is needed to integrate the two approaches. There have been three Argumentation in Multi-Agent Systems (ArgMAS) workshops already that have contained much work along these lines.

> ArgMAS 2004, Columbia University, NY
> http://homepages.inf.ed.ac.uk/irahwan/argmas/argmas04/
> Post-proceedings published by Springer

> ArgMAS 2005, Utrecht University, The Netherlands
> http://www.sci.brooklyn.cuny.edu/~parsons/events/argmas/argmas05/
> Post-proceedings published by Springer (Karunatillake,. Jennings, Rahwan and Norman, 2005)

> ArgMAS 2006, Future University, Japan
> http://homepages.inf.ed.ac.uk/irahwan/argmas/argmas06/
> Post-proceedings accepted for publication by Springer (Rahwan, Moraitis and Reed, 2005)

It is to be hoped that this new research will expand efforts to bring dialog theory in line with the implemented systems being designed and used in multi-agent technology.

What can dialog theory contribute to the development of ACL's? It has contributed quite a bit already, by introducing a new model of reasoning that is much more flexible than that represented in the traditional deductive logic. This new model is much more applicable to the kind of defeasible reasoning that is ubiquitous in computing. The forms of argument, represented by argumentation schemes, are (typically) defeasible, meaning that the probative weight of acceptance can be withdrawn by new evidence at a later point in an investigation. Argumentation is also conversational, or dialectical (dialog- based), meaning that it views an argument as a sequence of reasoning taking place between two parties, called a proponent and a respondent. Thus it views the concept of an argument as communicative in nature. The evaluation of argumentation is seen as pragmatic and contextual. An argument is judged to be good or not depending on how it was used for some purpose in a conversation that it is part of. An argument still retains a kind of local, premises and conclusion form (argumentation scheme), just as in

the traditional semantic approach to logic. But judging the argument as structurally correct or not now depends on how the argument was used at some given stage of an orderly goal-directed conversation. A lot therefore depends on the dialog structure of the kinds of conversations that are appropriate frameworks for different kinds of argument use.

The study of fallacies has shown that the six main types of dialog outlined in chapter 1 are especially useful for the analysis and evaluation of argumentation. Looking over the types of dialog in table 1.2, it has already been made evident that the middle four types are vitally important in multi-agent communication. And the importance of these types of dialog is easily accepted by anyone with even a little knowledge of multi-agent systems. But even those very familiar with multi-agent systems might wonder if the two remaining types of dialog are of any importance for ACL development. The persuasion type of dialog has not been in the forefront or played any central role in ACL's at their present state of development. The uses of the eristic type of dialog seem even more remote. The narrow focus on information-seeking communication is changing, however, as ACL's are applied to more sophisticated kinds of agent communication problems. The deliberation type of dialog is very important, as the focus on goal-directed practical reasoning in AI attests (Woodridge, 2002). Agents are supposed to carry out autonomous actions. Obviously therefore, deliberation will be a central type of dialog framework in which agents need to reason and communicate by planning together. The mass of literature on planning in AI abundantly testifies to the importance of deliberation as a type of dialog in which agents must take part. In order to carry out actions with other agents, negotiation is another type of dialog in which agents must take part. Negotiation can be viewed as a sequence of speech act moves, according to the negotiation protocol proposed by Parsons and Jennings (1997). The five speech acts in the protocol are proposal, critique, counter-proposal, explanation and meta-information. The first four speech acts are self-explanatory. Meta-information consists of any extra bit of information that is useful to guide the process along (Carbogim, Robertson and Lee, 2000, p. 137). This component suggests that an information-seeking dialog embedded within negotiation dialog could have a constructive function in helping the negotiation process along. In fact, Parsons and Jennings advocate the use of an argumentation approach to generate and evaluate the various proposals and other moves made in negotiations. In a very useful survey of argumentation-based models in multi-agent systems and other knowledge engineering systems, Carbogim, Robertson and Lee (2000, p. 138) also mention the work in Sierra, Jennings, Noriega and Parsons (1997), which applied multi-agent technology to negotiations in business. Their model of negotiation is based on an agent communication model that includes elements of persuasion. This innovation brings out the interesting feature of how persuasion dialog can be embedded in

negotiation. The inquiry is also a type of dialog that has some role in ACL's, although that role has not yet been investigated. The future direction in ACL research needs to expand in the direction of considering these types of dialog, and even how transitions from one type of dialog to another can occur in a sequence of argumentation. The theory of how such transitions occur is the subject of chapter 7.

Not only does multi-agent communication need to be expanded to consider more aspects of dialog theory, dialog theory itself needs to be expanded in order to more fully do justice to the notion that the proponent and the respondent in a dialog can act as agents. The entry point to this development is the notion in ACL's of the rational effect of a speech act. As described by Labrou, Finin and Peng (p. 49) the rational effect is a kind of forecasting device that is used in advance of a speech act in order to judge its potential effectiveness. The rational effect enables an agent to look ahead in a dialog with another agent, based on what the first agent knows about the second agent, so that the first agent can plan ahead in the dialog and know what speech act is best to perform next. In short, the rational effect can very nicely be viewed as a dialog strategy used by an agent communicating with another agent. Thus chapter 6 takes up the problem of how to model the participants in a dialog as agents with strategies.

Agents in critical argumentation

This chapter presents a new way to model the notion of fallacy in systems of formal dialectic. According to the pragmatic theory of fallacy in (Walton, 1995), a fallacy is a kind of sophistical tactic used by one party in a dialog to try to get the best of the other party in the dialog. But modeling the use of a sophistical tactic by a participant in a dialog requires some way of formally representing the notions that the participant is capable of carrying out a strategy of argumentation, and also has the capability for deception. Such capabilities are based on anticipation by one dialog partner of the expected moves of the other party, and also on the use of an organized attempt to exploit this anticipation by making tactical moves in advance. The problem is to figure out how the current systems of formal dialectic can be improved to take these capabilities into account and to model them in a way that is useful for analysis and evaluation of argumentation and fallacies. Another compelling reason for expanding dialog theory is that a participant in a dialog can be modeled as an agent with a strategy so that the enriched dialog theory provides a nice way of representing conversation policies in multi-agent systems. Instead of trying to formulate conversational policies in the confusing metaphysical psychology of the BDI model, these policies can now be understood, and the rationale behind the policy made sensible, by seeing them as based on different kinds of agent argumentation strategies.

The method proposed is to broaden the notion of a participant in a persuasion dialog by thinking of a participant as having several key properties of an agent in a deliberation type of dialog. An agent is an entity that not only has the capability for action, but that also has goals, and can modify its actions when it observes their immediate effects, in order to make them move closer to these goals. In a deliberation dialog, an agent can modify its plan of action once it sees its speech partner reacting in certain ways to its moves. An agent can also anticipate the expected plausible moves of its speech partner, and design its moves in advance, to fit in with these future expected moves. This proactive capability of agents was illustrated in the example of multi-agent communication presented by Huhns and Singh (1998), as outlined in chapter 1. In the current systems of formal dialectic, so far mainly used to model the persuasion (critical discussion) type of dialog, a

participant in a dialog is seen as a simple unanalyzed unit to which sets of statements called commitment sets are attached. A participant makes moves, has sets of commitments, and follows the dialog rules. But in the past there was no obvious way to model the notion of a participant capable of carrying out strategies of argumentation based on anticipation of the moves of the other party, and of sometimes even using calculated strategies of deception when making such strategic moves.

The notion of a participant moving in a proactive way like an intelligent agent in a conversation is, to some extent, implicit in Gricean conversation theory, and in the rational effects communication principles in ACL's. But the problem is that the notion of an agent currently in use in artificial intelligence, especially in the field of multi-agent systems, is quite complex, as shown in the last chapter. Due partly to the BDI framework in which agent communication tends to be viewed, the problem is that it is hard to know how to model conversation policies for agents in formal dialectic in a precise way. The approach taken in chapter 6 is to begin with certain key characteristics of an agent that show evidence of being useful for the study of fallacies, and examine how they could be modeled in a persuasion type of dialog. The next step taken is to add what is called a strategy to the commitment sets of both participants in a dialog. A strategy is seen as being a kind of plan, similar to the notion of plan widely used in artificial intelligence (Wilensky, 1983). It is also comparable to the notion of a partial strategy used by Hamblin (1987) in his analysis of the logic of imperatives. A strategy in a persuasion dialog is a device that a participant uses to anticipate the plausible future moves of the other party, and to weave these anticipated moves into her own projected sequences of argumentation.

1   The case of the critical discussion on euthanasia

Bob and Wilma are having a critical discussion on the subject of euthanasia. Wilma is opposed to euthanasia and Bob is for it. At one particular juncture of the argumentation, Bob is trying to argue for the conclusion that physician assisted suicide should be allowed in some cases. But he knows, from the previous argumentation, that Wilma's favorite argument against euthanasia is the slippery slope argument. According to this argument, if euthanasia is made a government policy, then it will be increasingly used to shorten lives, leading to a loss of respect for human life, and ultimately to organized mass killing of those deemed to be unfit or undesirable by those in power. At the present stage of the discussion, Bob is presenting a case in which a patient with advanced and incurable cancer had to suffer in agony while enduring chemotherapy and other painful treatments that Bob described as futile. This argument appears to be a powerful one, but Bob figures that later on, at some

strategic point, Wilma will bring out the main weapon in her arsenal, the slippery slope argument. But Bob has a strategy. From listening to Wilma's argumentation in the past exchange, he has noted that she makes a big issue of the claim that if a policy has not been tested out by being implemented on a wide scale, all kinds of unanticipated bad consequences could flow from the adoption of that policy. To counter this argument, Bob has his own counter-argument held in reserve. It is the argument that physician-assisted suicide has been tested out as a policy in the Netherlands, where it is perceived by the Dutch public as working well, and has not been subject to abuses of the kind that Wilma warns about.

The argumentation in this case can be reconstructed as follows. Bob knows that the slippery slope argument is a species of argumentation from consequences that depends on the following claim that is an instance of argumentation from consequences.

> If physician-assisted suicide is allowed as a practice, bad consequences will follow.
>
> We should not allow these bad consequences to occur.
>
> Therefore, we should not allow physician-assisted suicide as a practice.
>
> This argument is an instance of the argumentations scheme for argument from negative consequences.

Scheme for Argument from Negative Consequences (Walton, 1996, p. 76)

> Premise   If A is brought about, then bad consequences will occur.
>
> Conclusion   Therefore A should not be brought about.

There are three critical questions matching the argumentation scheme for argument from consequences (Walton, 1996, pp. 76–77).

> CQ1: How strong is the likelihood that these cited consequences may, will or must occur?
>
> CQ2: If A is brought about, will or might these consequences occur, and what evidence supports this claim?
>
> CQ3: Are there other consequences of the opposite value that should be taken into account?

Based on this scheme and its critical questions, Bob can know in advance that there are several possible ways he can appropriately critically question or attack this argument. One is that he can attack the first premise, by finding a case in which physician-assisted suicide was allowed as a practice, but bad consequences

did not follow. He can find a case, in other words, in which the antecedent of the conditional is true but the consequent is false. Since Bob knows of such a case, citing it could be a good move to use, if needed. Bob could also prepare a strategy in other ways. For example, he might anticipate that Wilma would use the counterargument that the Netherlands is different from North America, and therefore a policy that might work in the Netherlands might not work in North America. Bob might also plausibly guess that Wilma knows that the taking of recreational drugs is allowed in the Netherlands. It might well occur to her to argue that this shows how the Netherlands is different from North America, because (so she might plausibly argue), making drugs legal in North America is a policy that would never work. Wilma might be inclined to use the slippery slope argument again to support this claim. Once again Bob might try to guard against Wilma's line of argument by anticipating it, and thinking of some ways of arguing that the situation in the Netherlands is quite similar to that in North America, in some relevant respects. By thinking ahead in this way, and then reacting with counter-arguments that match Wilma's prior arguments, Bob will have devised strategies that are likely to be successful.

However, although the line of argument that Bob plans to use could be quite reasonable, the same kind of anticipatory plan of attack could be useful if you are attempting to commit a fallacy. For example, suppose that Wilma knows that Bob (or the audience of the critical discussion, if she is trying to convince them), tends to be conservative, and is susceptible to arguments that warn of the dangers of departing from a known and tested policy that has proven safe and acceptable in the past. She may even know that Bob is fearful of new things. She could take advantage of this knowledge by stressing, or even exaggerating the dangers of euthanasia as an untested policy that could have horrible results. She could argue, "It's the first step to Nazi concentration camps." The basic problem with this slippery slope argument is that it fails to support all of the required intervening steps between the first step and all the other next steps that lead towards the cited ultimate outcome. A slippery slope argument is hard to support adequately, because not only do enough of the missing steps have to be filled in to make the argument plausible, but it has to be shown how there is a gray area that the intervening steps lead towards, and during which there is no turning back (Govier, 1982; Walton, 1992). But instead of doing all the work of filling in all the missing parts of the argument that need to be supported by good evidence, Wilma may opt for the short cut of stressing only the horrible outcome, and appealing to Bob's fear of this outcome as a substitute for fulfilling the requirement of burden of proof. The deceptive tactic has several aspects. One is that Bob's support of euthanasia, because of Wilma's argument, now appears to make him somehow committed to or supporting Nazi concentration camps. Another aspect is that Nazi concentration camps are scary, and so Wilma's

argument raises all kinds of scary images that make euthanasia seem scary as well. But even though Wilma's slippery slope argumentation tactic is sophistical, meaning that it is based on a strategy to commit a fallacy, the same kind of planning by trying to anticipate how Bob will react is useful. It is useful in the sense that it is likely to make her argument to support her side of the case successful to persuade Bob to accept her claim in the dispute. Both sides are using argumentation strategies, and the use of such strategies is typical in a critical discussion.

This case represents a typical critical discussion in which a difference of opinions is supposed to be resolved or contested by allowing both sides to put forward the strongest arguments they can to support their respective conclusions. The stronger the arguments on both sides are, and the more they interact strongly with the arguments of the other side, the more revealing and insightful the discussion will be. In this case, Bob has a strategy, worked out in advance of the point where he will use it in the discussion. This strategy of anticipating what he guesses will be Wilma's main move could turn out to be quite a good one. If Bob were to try to counter Wilma's slippery slope argument even before she uses it, such unfortunate timing might even make it appear that he himself is putting forth the slippery slope argument. Far better to wait until she puts forward the argument, and then use the Netherlands counter-argument to attack the weak point in her version of the slippery slope argument. Then he can see how she presents the argument, and respond at that stage to what seem to be the weak points in her argumentation. He can also respond to whatever counter-arguments to his attacks that she might come up with. The tactic Bob uses is not only one of anticipation of Wilma's likely future counter-arguments to his arguments. It also involves reacting to each of her arguments individually once he actually hears them presented, and gets some idea of how she has reacted to his original arguments. The strategy Bob uses is a combination of anticipation and feedback, combining both foresight and testing out his own arguments to see what kind of reactions he is getting, once one of his arguments has been used.

Such uses of anticipatory argumentation strategies have not been very much studied in formal dialectic, even though informal fallacies do appear to be based on argumentation tactics of precisely this kind. But the literature in speech communication and rhetoric has paid more attention to such matters. Rhetoric is a different field from dialectic, but the two have significant overlap, particularly for one key reason. In some structures of dialog (like the critical discussion – see the classification of different types of dialog in Table 1.2 in chapter 1), each side in a dialog needs to use the strongest possible arguments to fulfill his or her aim in the dialog. As Houtlosser and van Eemeren (1998, p. 58) put the point, "rhetorical aims can play a legitimate part in a dialectical confrontation" because the discussants need to (and should) put forward arguments that will fulfill these aims. Thus the study of strategic

maneuvers in argumentation is not only appropriate for rhetoric, where they have traditionally been considered under the heading of "invention" of arguments. Some aspects of them are also quite important for dialectic to study. In particular, the use of anticipatory and feedback strategies to try to get the best of a speech partner in a dialog involving contestive argumentation is an important subject for study in the new dialectic. We could not analyze and evaluate argumentation in many common-place kinds of cases where arguments are put forward and fallacies committed, without having some grasp of how anticipatory strategies are used.

2   Fallacy and deception

Recent research on informal fallacies has moved more and more to seeing the fallacies not just as invalid arguments, using non-contextual models like deductive logic to model the fault. The current approach is to take the context of dialog into account. Fallacies are seen as tricky tactics that one party in dialog can use to unfairly get the best of the other party. Fallacies modeled in the new pragmatic analysis as moves in a dialog that interfere with the realization of a goal of collaborative dialog, or even block the dialog from moving forward appropriately towards its collective goal (Walton, 1995). But this pragmatic notion of fallacy involves more than just improper moves that violate the rules for a dialog. Fallacies are arguments, or moves in argumentation, that are improper in a dialog, but that have some appearance of looking to be appropriate, at least in some sort of dialog. In a word, fallacies are deceptive moves, or deceptive sequences of moves in a dialog. An explanation for how such deception occurs in some cases of fallacies has been ventured in (Walton, 1995). Fallacies often seem to be legitimate and reasonable moves in argumentation because there has been a deceptive shift from one type of conversations to another. Such deceptions can be effective argumentation tactics because at least one party to the conversation has expectations about the type of dialog that the argument is supposed to be part of. This expectation can be exploited by the other party to make a move in his argument seem reasonable and appropriate when really it is not. How could one construct normative models to represent the structure of the various types of collaborative conversation in order to systematically evaluate arguments in which fallacies may be committed? The answer according to Hamblin (1970; 1971) is to construct systems of formal dialectic. Hence in the next section, the basic components of the current systems of formal dialectic are outlined.

As noted above, fallacies are not only incorrect, or invalid arguments, but arguments, or other moves in argumentation that, as the saying goes, seem to be correct, or seem to be valid. But what does this "seeming" aspect amount to? It

doesn't have to mean that every instance of a fallacy is an intentional deception (Van Eemeren and Grootendorst, 1992). But the concept of fallacy does seem to be built on the notion of deception, or that of deceptive communication. A fallacy is a type of move in argumentation that is worth paying attention to because it can be used to powerfully deceive an arguer (Walton, 1995). It has the strong potential for deception. It represents the kind of case we should be worried about in many important contexts of the use of argumentation. In other words, there is an essential connection between the concept of fallacy and the concept of deception, even though not every instance of the committing of a fallacy has to be a case of the intent to deceive on the part of the perpetrator.

The lesson drawn is that the committing of a fallacy, like any deceptive communication, depends on expectations that both parties in a conversation have, in virtue of the type of goal-directed conversational exchange they are taking part in. Fallacies, like the deceptive tricks used by magicians to deceive an audience, are based on exploiting these expectations. Things are normally expected to go in a certain way by an audience watching a magic trick. And it is precisely this expectation that the magician exploits, in order to make things seem to be what they are not. But what exactly is deception? What are its components, and how does it work? Deception is very common in the behavior of animals. The study of animal communication can give us some insight into these questions.

Many animal communications are based on expectations about how another animal will act, once the other animal sees the actions of the first animal. Deceptive animal communication is especially interesting in this regard. A lizard, for example, may blow up a pouch in its neck so that face on, it appears much larger and more menacing to another animal approaching it. The lizard is projecting the message to the other animal, "Don't mess with me. I'm large and powerful, and there will be bad consequences for you if you try to attack me". Another interesting kind of case is the use of deception and camouflage by animals. A harmless snake may mimic the bright red color of a highly poisonous snake. Predators could easily eat the harmless snake, but will not do so if it appears to them that it is similar to the highly poisonous type of snake they usually try to avoid. A butterfly may have big eye-like patterns on its wings so that it looks like a fierce owl to passing predators. Or to take an even more common kind of case, a lizard may blend in with its surroundings by changing color to match its background wherever it moves. Such uses of deception by animals appear to involve communication of a sort. The kind of communication involved appears to be based on expectations by one animal concerning how another animal will plausibly react when confronted with a particular kind of situation. For example, the one animal appears to be acting on the assumption that the other animal would normally see it as prey, once it sees its characteristic outline or pattern of appearance. But if the prey animal

changes its color and blends in, the expectation is that the predator animal will not see the normal appearance that triggers its attack. In other words, the camouflage tactic can be seen as based on communication. And the basis of the communication is an expectation concerning what the other predator animal can plausibly be expected to do in a normal kind of situation that it regularly confronts. The deceptive camouflage tactic takes this action, based on normal expectation, and defeats it by changing or disrupting the normal pattern of appearances. Things are made to be something other than what they appear to be.

Another device used by animals to deceive potential predators is to play dead or to fake injuries. Ristau (1991) studied the behavior of the plover, when it feigns injury to lure a predator away from the nest. The plover drops a wing to the ground, as if it were broken, and then adopts an awkward flapping walk as it moves away from the nest. As soon as the predator gets close enough, the plover flies away. Results of experiments conducted by Ristau showed that plovers are sensitive to predator gaze and action, and respond more strongly to a dangerous predator. It is interesting to try to figure out what kind of thought process the predator may be going through when it engages in this kind of deception. One requirement surely is that the plover must anticipate the normal or expected reaction of the potential predator, and then try to exploit the anticipated reaction of the predator by sending out a message that will modify it. The predator wants food, supposedly, and would normally get food by taking the plover's eggs. That is the normal or expected action of the predator. But if the plover looks injured, such easy prey could be an even more evident and attractive source of food. Thus sending a message of injury not only anticipates that action, but re-directs it.

A comparable kind of tactic is the red herring fallacy, in which an arguer "diverts the attention of the reader or listener by changing the subject to a different but sometimes subtly related one "(Hurley, 2000, p. 131). The strategy of this sophistical tactic is to draw out a line of argumentation that leads away from the arguer's thesis that is supposed to be proved in the dialog. But this diverging line of the argumentation should nevertheless appear interesting to the audience, for the deception to work. The diverting line of argumentation should appear to be relevant to the issue that is supposedly to be resolved. The tactic in cases of such fallacies of irrelevance is similar to that used by the plover. The tactic of the deception works by distraction.

What is common to these instances of deception in animal communications is not that the one animal is intentionally trying to deceive the other. The snake does not intentionally equip itself with a red skin that makes it look poisonous to predators. It was born that way, and achieved its red skin (presumably) through evolution. What is common to these instances of deception is that the message sent out is deceptive because it anticipates and exploits the expected reaction of the mes-

sage recipient. The snake predator sees a red skin color which it takes to be poisonous, or to indicate deadly poison. That would normally be the right reaction, on the basis of normal expectations. The predator may not be sure that this snake is poisonous. But safety suggests acting on that presumption. What is evident is a deployment of plausible inference, triggered by a certain appearance which suggests something – in this case danger. The communication is between two animals – the message sender and the receiver, or potential predator or attacker. The message is deceptive because the receiver takes it in a kind of routine or normal way, which is in this case however, misleading. Animal deception shows a complex dynamic that is characteristic of deceptive communication. The message sender exploits the normal and expected reaction of the message receiver, getting the receiver to draw a wrong conclusion.

Now the problem is to ask how this characteristic dynamic of deception could be modeled in a formal dialectical system. Dialectical systems are not empirical models, but normative models of argumentation. We shouldn't expect them to model the empirical aspects of deception. That is a job for the empirical social scientists. But once again, certain key aspects of deception do need to be modeled if justice is to be done to the pragmatic notion that a fallacy is a deceptive tactic of argumentation. The formal dialectical system should model certain key communicative aspects of the kind of reasoning used when a fallacy is committed because a message receiver draws a wrong conclusion. Basically, the model needs to grasp how the message sent out transmits cues to the respondent so that, given the normal expectations of the respondent, he draws a plausible inference, and yet the inference fails to be one that leads to the right conclusion. It is this dialectical dynamic that the dialectical structure needs to model.

3   Current systems of formal dialectic

As shown in chapter 1, the systems of formal dialectic that have been put forward as normative models that are useful for the study of fallacies (Hamblin, 1970; 1971) are dialog structures in which two parties take turns making moves. According to Hamblin's theory, as outlined in chapter 1, a dialectical system is a regulated dialog with two (in the simplest case) participants. The structure of such a dialog has three basic components. First, a dialog has two participants, a proponent and a respondent, often called 'White' and 'Black' respectively by Hambin. Second, there is a set of moves (typically questions and replies) made by the two participants. Third, the moves are made in an orderly sequence. By convention, the proponent makes the first move, the respondent makes the next move, and then they proceed to take turns, in an orderly manner. Formally, any dialog can be modeled as a se-

quence of such moves, by listing all the moves in order of occurrence. Each member of such a sequence is defined by Hamblin (1971, p. 130) as a triple, ⟨*n, p, l*⟩. *n* is a number representing the length of the dialog (the number of moves so far). *p* is a participant, and *l* is a locution. As the participants take turns making moves, of a kind allowed by the rules, the dialog unfolds. The rules specify the content of what they say, relative to the context of dialog and the prior moves in a dialog. Hamblin illustrates how a dialectical system works by giving a simple example of what he calls a Why-Because System with Questions (1970, p. 265). The following kinds of moves are illustrated.

1.   The making of assertions, in the form 'Statement *S*'.
2.   The asking of questions, in the form '*S*?', for example.
3.   The retracting of commitments.
4.   The request for justification, in the form 'Why *S*?'.
5.   The making of a resolution request, in the form, 'Resolve *S,T*'.

A resolution request would be appropriate when two commitments conflict, as illustrated in the profile of dialog represented in the tableau in figure 1.1 of chapter 1. In this kind of situation, the resolution request requires the proponent to choose between the two statements she has committed herself to. The rules determine not only what kinds of moves can be made, but also what happens as the result of a certain kind of move. For example, whenever a participant asserts a statement *S*, then *S* automatically is inserted into her commitment set. Or if a participant retracts *S*, then *S* is removed from his commitment set. Hamblin envisioned different kinds of dialogs with different kinds of rules for each. But he made no systematic attempt to classify the different types of dialog. Judging from his various remarks on the various systems of formal dialectic he constructed however, it would probably be fair to say that the central type of dialog he had in mind would correspond roughly to what is now called persuasion dialog.

Indeed, from a point of view of studying the fallacies, the most central and significant type of dialog to consider as a beginning point is the persuasion dialog. In a persuasion dialog, the proponent has a designated thesis to be proved by rational argumentation. The respondent either has the task of casting doubt on that thesis by raising questions about it, or the task of proving the opposite of the proponent's thesis. In a persuasion dialog, each party takes the initial concessions of the other as premises, and then by a series of steps, tries to use these statements as premises in rational arguments designed to persuade the other party. What does "persuasion" mean in this context? To successfully persuade your respondent, your aim is to fulfill your task in the dialog using only premises that your respondent is either already committed to, or can be gotten to accept by steps of rational argumentation. In other words, in a persuasion dialog, empathy is all-important. Your

arguments must always be based on the commitments of the other party. Thus this notion of persuasion is really one of rational persuasion. Persuasion of a party in a dialog is successfully achieved only by using a structurally correct sequence of argumentation steps such that each premise utilized is a commitment of the other party. Different kinds of rules for different kinds of persuasion dialogs have been constructed in (Walton and Krabbe, 1995). Other types of dialog have goals that are different from that of the persuasion dialog. For example, the goal of negotiation is to "make a deal" to resolve a conflict of interests. In contrast, the goal of a persuasion dialog, as indicated above, is to resolve a conflict of opinions. Hence the kinds of arguments appropriately used in the one type of dialog, could be out of place, or even fallacious, in the other type of dialog.

There is a method of profiles of dialog outlined in chapter 4 can be used to formally model any given dialog sequence of the kind that occurs in a Hamblin dialectical system. As shown in figure 1.1 in chapter 1, the tableau lists the sequence of all of one participant's moves in a left column, and lists each of the other party's moves displayed in a right column. As shown in chapter 1, the tableau is one method of representing a profile of dialog. A Hamblin-style tableau can be used, for example, to represent the dialog sequence in profile 2 (chapter 4, section 2). Each pair of moves is numbered. The tableau displays how each move relates to the prior and subsequent moves in the dialog. Following Hamblin's notation, in table 6.1 below the letters S, T, U,…, stand for statements.

Table 6.1 A Hamblin-style tableau

Proponent	Respondent
1. Why S?	Because T is true, and T implies S.
2. Why should I accept T?	Because U is true, and U implies T.
3. I do accept U.	Do you accept T?
4. Yes.	Do you accept S?
5. No.	But you must, because T implies S.

Table 6.1 exhibits a profile of dialog in which the proponent begins by asking the respondent to justify S. The respondent replies by putting forward an argument. The proponent, however, declines to accept one of the premises, and requests the respondent to give her a reason to accept it. The respondent replies by putting forward another argument. The proponent indicates that she does accept the premise this argument is based on. Then the respondent asks the proponent whether she now accepts S, the statement she originally doubted at the first move. At move 5, she replies 'No.' The respondent then points out that she had accepted T at a prior move. He uses the additional assertion that T implies S to show that she

must accept *S* now, unless she wishes to challenge the statement that *T* implies *S*. Table 6.1 represents a typical sequence of argumentation in a Hamblin dialog. One party is using inferences to try to get the other party to accept some particular proposition, and the other party is expressing doubts. This pattern is typical of the many examples and profiles of dialog considered by Hamblin. It is for this reason that it was suggested above that Hamblin probably had something like the persuasion dialog in mind as his main model, although he never made any definitive statements about the global goals of types of dialog.

Subsequent developments in dialog theory followed Hamblin's model in general outline. Each system has moves, rules, and participants. But different systems went different ways. The Hintikka systems are based around the notion of an information-seeking type of dialog. The systems in Barth and Krabbe are mainly based around the central notion of the persuasion dialog, but also acknowledge the inquiry type of dialog. The Kripke model for intuitionistic logic can be seen as a dialog system, but it appears to represent the inquiry type of dialog. Now the various types of dialog have been classified, as shown in chapter 1, these various formal dialectical systems can be put in a broader perspective. It can be seen how and why each system is different, because it represents a different type of dialog with a different goal. And yet each system has the same main components. In many ways, the Hamblin outline of formal dialectic displays these main components most clearly and systematically. Looking at the Hamblin systems of formal dialectic, you can get a good idea of what is there, and how what is there is fairly minimal. It is a basic structure that can be developed and enriched in various ways. It sets out a clear foundational basis that can be built on in various ways, depending on what you want to do with it.

Four basic and simple formal systems of dialog presented in (Walton, 1984) model argumentation much as in the Hamblin style. They are meant to represent systems for persuasion dialog in which one party has a designated thesis to be proven and tries to use rational argumentation based on the other party's commitments to try to get him to come to accept this thesis. One of these systems, called CB, also uses the notion of immediate consequence due to Mackenzie (1981). A statement T is said to be an immediate consequence of a statement S if T follows from S by means of applying one rule of inference (like *modus ponens*) to S in one step. There is assumed to be a non-empty set of rules of inference in the game. For example, *modus ponens* could be a rule of inference, but rules of inference that allow infinite repetitions like 'S, therefore S ∨ T' are not included. Following Mackenzie (1981), a statement T is said to be an immediate consequence of a set of statements S_0, S_1,..., S_n if and only if 'S_0, S_1,..., S_n, therefore T' is a substitution instance of an inference rule in the dialog system. A statement T is said to be a consequence, whether it be immediate or mediate, of a set of statements S_0, S_1,..., S_n if

and only if T is derived by a finite number of immediate consequence steps from immediate consequences of $S_0, S_1,..., S_n$. CB is a simple dialog system that does not allow for the more complex kinds of moves, like certain kinds of questions allowed by Hamblin, or commitments to challenges of the kind allowed in Mackenzie's systems. CB was designed to be a basic system that can provide a minimal platform that can be extended to modeling varieties of more complex dialogs.

In the version of the rules of CB given below, the only changes from (Walton, 1984, pp. 133–135) are the names of the rules (indicated below as CBLR1 and so forth). The locution rules, commitment rules and dialog rules are easily recognizable as being of the Hamblin sort.

Locution rules

CBLR1. Statements: Statement letters, S, T, U,..., are permissible locutions, and truth-functional compounds of statement-letters.

CBLR2. Withdrawals: 'No commitment S' is the locution for withdrawal (retraction) of a statement.

CBLR3. Questions: The question 'S?' asks 'Is it the case that S is true?'

CBLR4. Challenges: The challenge 'Why S?' requests some statement that can serve as a basis in (a possibly defeasible) proof for S.

Commitment rules

CBCR1. After a player makes a statement, S, it is included in his commitment store.

CBCR2. After the withdrawal of S, the statement S is deleted from the speaker's commitment store.

CBCR3. 'Why S?' places S in the hearer's commitment store unless it is already there or unless the hearer immediately retracts his commitment to S.

CBCR4. Every statement that is shown by the speaker to be an immediate consequence of statements that are commitments of the hearer then becomes a commitment of the hearer's and is included in his commitment store.

CBCR5. No commitment may be withdrawn by the hearer that is shown by the speaker to be an immediate consequence of statements that are previous commitments of the hearer.

Dialog rules

CBDR1. Each speaker takes his turn to move by advancing one locution at each turn. A no-commitment locution, however, may accompany a why-locution as one turn.

>CBDR2. A question 'S?' must be followed by (i) a statement 'S', (ii) a statement 'Not-S', or (iii) 'No commitment S'.
>
>CBDR3. 'Why S?' must be followed by (i) 'No commitment S' or (ii) some statement 'T', where S is a consequence of T.

Several of these rules have already been discussed in sections 9 and 10 of chapter 2, where alternatives to them were considered. Here we should not be concerned with the issue of whether one particular rule or some possible alternative is best. The rules above are simply set out as minimal examples of the kinds of rules needed.

CB above exhibits all the main components of a dialectical system, according to the current development of dialectical structures. There are two participants, and each makes moves in the dialog according to certain rules. The structure is fairly simple, but has normative bite, as applied to modeling argumentation in any real case, because each type of dialog has a goal. It clearly is an excellent tool for the study of fallacies. But the concept of fallacy is actually quite deep and complex in certain ways that make it problematic to model (at least very deeply) in a simple Hamblin dialectical structure. The basic problem is that there is no obvious way to model deception in such a structure, in a way that would be useful to analyze the pragmatic concept of fallacy. How could such a structure be enriched to model the use of deception in a sequence of argumentation? What is required is for the arguer to be aware of how her dialog partner can reasonably be expected to react to a move, or sequence of moves, before the move is made. What is required is to build the notion of anticipation into the structure of the argumentation tactic used when such a sequence of moves is put forward in a dialog.

A new system called ASD (Reed and Walton, 2007) extends CB by adding some new rules that enable argumentation schemes to be added to the inferential structure of CBV. One of these new rules is a new locution move that allows for critical attacks to be posed by one party against an argument put forward by the other (p. 6).

>ASDLR5. The attack 'Pose C' poses the critical question C associated with an argumentation scheme.

Another rule of ASD is the following dialog rule (p. 5), paraphrased below.

>ASDDR4. After a statement T has been offered in response to a challenge locution, 'Why S?', then if (S,T) is a substitution instance of some argumentation scheme recognized in the dialog, the locution 'pose C' is a legal move, where C is a critical question matching the scheme.

This rule allows not only for the addition of argumentation schemes to the deductive rules that were allowed in CB, but it also introduces a locution 'pose C' that allows for the asking of critical questions as part of the dialog.

ASD broadens the usefulness of CB considerably by allowing for the introduction of argumentation schemes that represent defeasible types of arguments, like argument from expert opinion and practical reasoning. It also allows us to model two different kinds of implicit premises in an argumentation scheme. An assumption is an implicit premise of an argumentation scheme that is assumed to be true, even though it may not be explicitly stated among the ordinary premises of the scheme. An exception is an implicit premise of an argumentation scheme that is assumed to be false, but later in the dialog, if it is found to be true, would defeat the argument. ASD handles these two kinds of implicit assumptions of argumentation schemes in different ways by postulating a different burden of proof for each of them. After a critical question, 'Pose C', if C is an assumption, the respondent must commit to either C, or to its negation, or state that he has no commitment to C. After a critical question, 'Pose C', if C is an exception, the respondent also has the option of shifting the burden of proof back to the proponent by asking the question, 'Why not-C?.' The capability for distinguishing in this way between assumptions and exceptions as implicit premises of an argumentation scheme is fundamental to the capability of dialog theory to deal with problems of implicit commitment.

4   Implicit commitment and Gricean implicature

Reacting to another party's argument in a dialog is typically based on interpreting that party's argument by being able to realize that the argument is based on unstated premises as assumptions. Hamblin (1970; 1971) required that the commitment-store of each participant in a dialog be a set of public statements, for example a set of sentences written on a blackboard in view of all the dialog participants. But there is an extension of CB called CBV (Walton, 1984, pp. 252–254) that allows implicit commitments as well as explicit commitments to be included in participants' commitment sets. Implicit commitments called "veiled" or "dark" – are not visible to either party, unlike explicit commitments that are on public view and known to all parties in the dialog. In CBV, implicit commitments of a party are revealed if the party having an implicit commitment makes a move to conceal it. For example, suppose the party denies he is committed to a particular proposition, but it is actually among his implicit commitments. In CBV he has to resolve the inconsistency by either retracting the implicit commitment or going on record as accepting it as an explicit commitment.

CBV is obtained by adding implicit commitment sets to CB, and by adding the following rule governing commitments sets.

> CBCR6. If a player states 'No commitment S' and S is in the implicit set of his commitments, then S is immediately transferred into the explicit set of his commitments.

But clearly CBCR6 is not the only way implicit commitments are added to incomplete arguments that are put forward in discourse in everyday conversational argumentation. Many implicit premises or conclusions in such arguments are based on other factors of kinds that will be examined in chapter 7, where we turn to the task of analyzing such arguments. At this point, it is important to consider gaps in a conversation created by implicit assumptions can be filled by recognizing that there are rules for collaborative conversation that each party expects the other will follow.

The role played by expectations when one party in a conversation draws conclusions based on the prior moves of the other party was very well revealed by the classic paper of Grice (1975) on conversational rules and policies. Grice observed that when a speaker who is a participant in a collaborative conversation "flouts" (overtly breaks) a rule (which Grice called a conversational maxim), the other party may judge that the first party has broken the rule intentionally. The reason is that the second party may be well aware that the first party is aware of the rules in question. The first party will also be aware that the second party will know what is going on. Therefore the second party may draw the conclusion that the speaker is communicating some message by that action. The whole process is based on expectations of how each party can expect the other party to normally move in a conversation. When such an expectation is violated, a conclusion is drawn about how to explain the violation, and what to conclude from it. The process sounds convoluted. Some examples will help bring out how expectations are involved.

Grice (1975, p. 72) considered a case in which, during the course of a conversation, someone makes an irrelevant remark. For example, suppose the discussion is about some other person, Ms. Y, known to the discussants, when suddenly one of them, Ms. X, says, "The weather has been quite delightful this summer, hasn't it?" From a viewpoint of the current type of dialog theory, all one can say here is that this statement is irrelevant, and does not match up with the prior sequence of exchanges in the dialog. But as Grice pointed out, the participants in such a conversation might draw quite a significant further conclusion. They might all be aware that Ms. X has changed the conversation to a different topic on purpose, showing that she was embarrassed about the prior discussion, or at any rate, didn't think it should be continued any further. To indicate this, Ms. X changed the topic to something completely unrelated to the prior discussion. Perhaps she felt that the others were delving too deeply into the personal affairs of Ms. Y, and she was

embarrassed by this, and felt it was improper. However, it would be impolite to just say this, directly. So she made the remark about the weather. By what Grice calls implicature, the others in the conversation figure out what is going on. They conclude that Ms. X is not just acting at random, but is sending them the message that the previous conversation should come to an end, or that she no longer wants to take part in it.

But now the problem is – how do the other participants in the conversation figure this out? All Ms. X did was to violate a rule of the dialog. Instead of continuing with the sequence in progress, she switched abruptly to some entirely different subject with no apparent relevance to the preceding conversation. From a point of view of the current dialog theory, all that is determined is that the remark of Ms. X violated a conversational rule. But from the Gricean point of view, more needs to be said about the case, and a further conclusion can be drawn. The reason is that Ms. X "flouted" the maxim of relevance, to use Grice's term. It looked like she intentionally and knowingly violated the rule of relevance. Why? The conclusion drawn by the others, as a best explanation of what happened, is that Ms. X was unhappy with the direction the conversation was taking, and she no longer wanted to take part in it. Also, she wanted to convey this fact to the others, but in an indirect way that would not be too impolite. Just not answering, for example, or saying that the others were discussing personal affairs of Ms. Y, that are none of their business, would be impolite.

Another example Grice (1975, p. 71) gives is a flouting of the conversational postulate he calls the first maxim of quantity, which tells a participant to make her contribution as informative as is required for the exchange, but not more informative than is required. In the example, a professor is writing a letter of reference for a student who is applying for a teaching job in philosophy. The professor writes in the letter only that the candidate's grasp of English is excellent and that his class attendance has been regular. How would someone who is thinking of hiring the candidate interpret such a letter? Grice commented (p. 71) that she would reason that since the student is this professor's pupil, he cannot be failing to furnish more information because he does not possess it. Therefore, he must be "wishing to impart information which he is reluctant to write down." The conclusion drawn is that the professor, by conversational implicature, is communicating to the reader of the letter the conclusion that the candidate is no good at philosophy.

So what is going on in these kinds of cases? What is going on is that as a conversation proceeds, the participants will have expectations, within certain boundaries, about how the next speaker will reply, or contribute to the conversation. If you ask a question, for example, you expect you will get either an answer, or at least a reply that indicates why no answer is possible, or that gives some kind of helpful advice on how an answer might be sought, or something of that sort. You expect

some kind of helpful reply that at least addresses the request made by the question. You don't expect a rude noise, a remark about the weather, or a statement like 'Love is blue', unless those replies relate to the prior sequence of conversation, or somehow contribute to the discussion. So if you get a reply like that, you might seek some reason or explanation for it. What is happening in such a case then, generally speaking, is that any participant in a conversation will have expectations about how the other party will move next, even before that other party makes any move at all. These expectations are an important aspect of the conversation and how it works. A deviation may trigger, by implicature, a suspected reason for the unexpected move, which will in turn enable the first party to draw a conclusion about what the first person is really saying, by making such an unexpected move.

Grice's observation about how language is used to convey a message indirectly in collaborative conversations is based on the assumption that speakers and hearers in such conversations have expectations on how the other party will normally respond at a given point in a conversation. When a participant acts in an unexpected way, conclusions will be drawn from that deviation from the expected norm. In the kind of dialog theory developed so far, there are rules for different types of moves. The rules determine whether a given move made by a participant in a given sequence of dialog is permissible or not. But if a participant reacts in an unexpected way, by breaking the rules and producing some odd or unexpected reply for example, all that is shown is that a rule has been violated. But Grice clearly wanted to go beyond that, and to work in a framework in which some conclusions are drawn in such a case.

5   Adding speech acts and agents to formal dialectic

Implicature works by expectation and feedback. It is these additional dimensions that they add to the current dialog theory structure. The participants in the dialog are no longer just units that can carry commitments. They now have expectations, and when they see an unexpected consequence of some action they have carried out, they draw conclusions, and perhaps try to correct their next actions accordingly. So Gricean conversational implicature does, at least to some extent, require seeing the participants in a dialog as agents that have expectations and are capable of actions, at least speech acts, of seeing the effects of these actions, and of reacting by drawing conclusions when these effects are unexpected. In order to carry out collaborative actions, an agent will need to communicate with other agents. But speech actions are just another kind of action that the agent carries out. When an agent puts forward a speech act in a dialog, she can plan the move, based on how she expects the other party in the dialog to react to it. An agent who carries out a

speech act in a dialog can also get feedback by seeing how the other agent reacts to this speech act. The other agent may react to it, indicating, for example, that he understands what was said, or that he will agree to the request contained in the speech act. Or if the original speech act was a question, the other agent may give a reply to the question. In one way, speech acts are just another kind of action by an agent, and don't seem to pose any special problems. But in some ways, speech acts are special kinds of actions for an agent. A speech act, like the asking of a question or the making of a promise, is a conventional type of act which, when made in the presence of another agent, requires the other agent to reply in a certain kind of way (Van Eemeren and Grootendorst, 1984; 1992). Speech acts, like all acts carried out by agents as part of a collaborative group effort, are based on expectations that one agent has with respect to how another agent may be presumed to act in response. The hypothesis suggested is that the moves in a dialectical system should be seen as speech acts. This hypothesis has already been proposed by the Amsterdam School (Van Eemeren and Grootendorst, 1984; 1992). But what is also required is that each participant in a dialog be seen as an agent, an entity with the capability for action, and for forming a plan of action, in anticipation of future expected events, and actions by other parties.

A good example is the speech act of proposing, a central type of move in any deliberation dialog. Proposing lies between a directive, a speech act in which the speaker tries to get the hearer to carry out a course of action, and a commissive, a speech act that commits the speaker to a course of action (Searle, 1969). In the speech act of proposing there is a choice between alternative courses of action, and one of the agents picks out one of these actions as something the two of them can commit to as a basis for collaboration. The following table, following the general format of the table in (Aakhus, 2006, p. 406), contrasts the speech act of making a proposal with two other classic speech acts analyzed by Searle. To distinguish proposing from the classic speech acts of requesting and promising, the new version of the speech act of proposing (Walton, 2006a) is presented in the column between them.

When two parties are acting together to solve a problem one of them may put forward a proposal to get the other to consider some proposition as worthy even if he had doubts about it, or might have dismissed or overlooked it previously. To achieve success in this speech act, the one who puts forward the proposal must support it by drawing out the doubts and objections against it, thus paving the way for its acceptance in the deliberation.

Table 5.1 Felicity conditions for requesting, proposing and promising

Act	Request (Searle, 1969)	Propose	Promise (Searle, 1969)
Propositional Content	Future act A of H.	Future act A of H + S.	Future act A of S.
Preparatory Condition	H is able to do A. S believes H is able to do A. It is not obvious to both S and H that H will do A in the normal course of events of his own accord.	H and S are able to contribute to the accomplishment of A. It is not obvious to both S and H that either S or H can do A of their own accord in the normal course of events. That A will leave neither S nor H worse off than not doing A.	S is able to do A. S believes S is able to do A. It is not obvious to both S and H that S will do A in the normal course of events of his own accord.
Sincerity Condition	S wants H to do A	S believes A will mutually benefit H and S or that if it benefits S it will leave H no worse off.	S intends that in uttering to do A he is under the obligation to do A.
Essential	Counts as an attempt to get H to do A.	Counts as an attempt to enlist H in mutually bringing about A.	Counts as an attempt to commit S to do A.

An important tool closely allied with deliberation is that of plan recognition technology, as explained in chapter 1, section 6. In the example of the empty car with a missing tire parked on the highway (Carberry, 1990, p. 17), the passing motorist inferred by plan recognition that the man sighted was the driver of the car. When she saw the children with him, she also inferred that the man had thought it unsafe to leave the children alone in the car. By observing the external facts of the situation, the motorist drew conclusions about what she took to be the goals of the man rolling the tire, and about his actions as presumed means to achieving these goals.

Plan recognition is based on simulative reasoning, in which one agent understands the thinking of another agent by being able to go through the same sequence of reasoning. Simulative reasoning of this kind tends not to be deductive, as noted in chapter 1. The one agent is confronted with a set of facts concerning the external behavior of the other agent. By finding a best explanation of these facts, the one agent can hypothesize the plan of the other agent. Thus plan recognition works by abductive reasoning. But the problem is how abductive reasoning of this sort can be verified, or based on good evidence, since the one agent can never directly observe what is in the other agent's internal plan.

How then should we enrich the current formal systems of dialectic to take agents, and agent-based technologies like plan recognition into account? What concept of agent is best? As steps toward solving the problem, in the next section an analysis is given of (1) what agents are, (2) how agents act, (3) how agents can act with other agents collaboratively, (4) how agents engage in communication with other agents, and (5) generally, how agents can engage in deceptive communications with each other. Shown to be vitally important are the expectations one agent has about how another agent will be likely to react when the two of them are communicating collaboratively with each other. An agent, it is argued, needs to be seen not only as an entity that acts, but as an entity that anticipates how another agent can be expected to act or react in a normal kind of situation that both are familiar with. This element of expectation turns out to be an extremely important characteristic of an agent for understanding multi-agent communication.

Introducing the concept of an agent into formal dialectic would involve some important complications. In the Hamblin systems (1970; 1971), in the systems of Barth and Krabbe (1984), and in newer systems deriving from these two traditions (Mackenzie, 1990; Walton and Krabbe, 1995), a participant was seen as a neutral entity that mainly served as a repository for commitments. Not much was made of the concept of a participant. The participant is just a designated individual that can make moves, like making an assertion, and that then becomes committed to propositions on the basis of having made a certain type of move. But if the participant is seen as an agent, much more begins to be involved. The agent can not only carry out actions by making moves in a dialog. An agent, as defined above in section 1, also has the capability of seeing the results of its moves on the other agent who is taking part in the dialog. The agent has feedback capabilities. Therefore, the agent will have expectations – that is, the agent will try to anticipate how the other party in the dialog is likely to respond to his move. An agent is a much more robust entity than a "participant" in one of the Hamblin dialectical systems.

Agents can base their arguments on anticipating how the other party in a dialog is likely to react to an argument – for example, by anticipating objections, and framing the argument to counter such objections in advance. This type of anticipatory device of argumentation is called *prolepsis* in classical rhetoric. Agents are capable of planning in advance to deal preemptively with kinds of problems or situations that they may expect to arise in the future. In other words, an agent has the capability of making up a strategy. In a conversation, an agent can make up some carefully plotted strategy to deal with the anticipated arguments or objections of her speech partner. The other agent in the conversation might also anticipate such a strategic line of argumentation, and devise a defensive strategy designed to counter it. These notions make the concept of an agent conversation

richer than that of a conversation between mere participants in a dialog in the older sense.

6   What characteristics of an agent are needed?

An agent is an entity that has goals, and that takes actions to try to fulfill these goals (Wooldridge and Jennings, 1995). These are the two primary characteristics of an agent. But an agent has two other defining characteristics as well (Franklin and Graesser, 1996). An agent has the capability of receiving information about what is happening in its immediate environment. In particular, it has the capability of seeing some of the immediate consequences of its own actions. This capability gives rise to a fourth characteristic called feedback. An agent, once it sees consequences of its actions, can change its subsequent actions in light of these perceived consequences, and so "correct" its actions by steering them more towards the goal it is aiming at. The four capabilities are the central defining characteristics of what an agent is.

What kind of reasoning do agents engage in? The answer is that agents characteristically engage in what is called practical reasoning. Practical reasoning is a kind of goal-directed reasoning that culminates in an action, or at least a decision that some action is a prudent line of action to take in a given situation (Walton, 1990). Practical reasoning is based on the information that an agent has about its situation. It is also a goal-directed type of reasoning. Practical reasoning is a chaining together of inferences that have the form of the Aristotelian practical syllogism (Clarke, 1985). In such an inference, the agent uses the argumentation scheme for practical reasoning. The scheme below is a simplified version of the one presented in chapter 4, section 1.

Simplified argumentation scheme for practical reasoning

> *G* is my goal.
>
> To bring about *G*, I need to bring about *A*.
>
> Therefore I need to bring about *A*.

The argumentation scheme above represents only the necessary condition type of inference in practical reasoning. The sufficient condition inference, as indicated in chapter 4, section 1, has essentially the same structure except that the second premise cites a sufficient condition (Clarke, 1985, pp. 43–63). In practical reasoning, the necessary and sufficient schemata are combined in longer sequences that include subinferences made up of both types of inference. To see how practical reasoning should be evaluated in any given case, the dialog structure of the argu-

mentation needs to be taken into account. The proponent puts her argument forward based on the argumentation scheme for practical reasoning. The respondent can then ask critical questions of the kind cited in chapter 4, section 1.

The 'need to' in the conclusion of the argumentation scheme represents a so-called prudential ought. It means that the agent has a reason for carrying out action A assuming that the two premises hold. It is assumed that the agent's goal really is G, and that B really is an action required to achieve G. Practical reasoning is defeasible, and not conclusive, however, as there may be other significant factors involved in a case, as indicated by the critical questions. For example, there may be more than one way to carry out goal G. One way may be better than another, meaning it may be more efficient, less costly, or otherwise have less negative consequences, from the agent's point of view, than another way being considered. Another complication is that an agent may have several goals, that need to be considered together. Carrying out one goal, in the given situation, may even conflict with carrying out another goal, and such a conflict needs to be resolved before a plan for action can be decided upon or carried out (Chu-Carroll and Carberry, 1995). Yet another complication is that fulfilling a goal may require carrying out prior actions, in order to carry out some needed action. So there may be a whole sequence of actions that the agent must carry out in order to achieve a goal, and not just a single action. Finally, *CQ4* indicates the complication that the agent may find out that the action being contemplated is simply not possible to carry out. According to Chu-Carroll and Carberry (1995, p. 115), this factor has been recognized in the literature on multi-agent planning by Martha Pollack, who argues that a plan can be invalid if one of its actions is "infeasible", meaning that the action cannot be carried out by the agent.

These various complications raise the whole question of what an action is. Agent communication languages presuppose some theory of action on which they can be based, as shown in chapter 1. Several leading theories have been proposed in the field called philosophy of action. Whatever an action is, it is something that is normally part of a connected sequence of actions called a chain of actions by Goldman (1970). Familiar sequences of actions of recognizable kinds are called routines (Segerberg, 1985). For example, take the apparently simple action of cutting wood. This action may involve quite a complex sequence of subactions. First I start the chainsaw by turning on the switch and then pulling the cord. Once the saw is running, I set the log in place. Then I pull the trigger on the saw and bring the chain in contact with the side of the log. Then the saw starts to bite into the log as the cutting process begins. The action as a whole sequence may represent a standard routine for me that I am very familiar with. I perform it in a practiced way, and take many safety precautions along the way. This practiced way of carrying out such an action is called a routine.

When an agent engages in practical reasoning, one of the critical questions that should be considered is whether or not it is possible to carry out the goal it has in mind. The assessment of whether the goal is possible will be based on the agent's stock of routines it has available. Some of the routines the agent is familiar with may get it part way to the goal. The question may be whether the missing bits of the required sequence of actions can be filled in by actions the agent can carry out. The assessment of this question in a given case is yet another complication in practical reasoning that makes the reasoning defeasible. The argumentation scheme for practical reasoning and the set of matching critical questions are important tools for a problem that occurs in plan recognition. One of the problems of analyzing an inferred plan occurs in the kind of case where it appears that the plan is not feasible because something indicates that the action indicated is not possible. Carberry (1990, p. 251) offers the example of a dialog where the proponent says to the respondent, "I want to get a soda. May I borrow 75 cents for the machine downstairs?" But suppose the respondent knows that the machine is empty. Thus even though the agent's plan is generally a good one for achieving this goal, it will not work in this case. The appropriate response would be for the respondent to say that he would be willing to lend the proponent the 75 cents, but that the machine is empty, and that perhaps she should try another machine in the next building. In this kind of example, the respondent needs to recognize a plan that is impractical because the indicated action is not possible to carry out, given the circumstances of the case. In general, plan recognition technology can be aided by understanding how apparent weaknesses in a plan evident to one agent can be taken into account when that agent is trying to recognize the plan of another agent.

When an agent attempts to carry out a goal in a systematic way by trying to anticipate what is likely to happen in the future, and take steps accordingly, the activity of planning should take this future orientation into account (Bratman, 1987). Plans involve not only goals, but attempts to carry out actions that will presumably lead at some future time to the fulfillment of the goal by actions that an agent can begin to carry out now. What is important then in a practical plan are projections on how events are likely to go in the future. Particularly important is the attempt to guess what the likely consequences of one's actions will be, based on expectations about how things normally go in domains that one is familiar with. The problem is that nobody knows for sure what will, or even might happen in the future, in a particular case. Trying to look into the future requires making up plausible hypotheses about likely outcomes of events that are at least partly uncertain. Planning also frequently involves coordinating the actions of one agent with the actual or predicted actions of another agent. In literature on planning in artificial intelligence (Schank and Abelson, 1977), a plan is a connected sequence of action steps contemplated by an agent who is trying to carry out a goal. The sequence is

essentially one of practical reasoning as applied to a given situation as the agent sees it, but an agent's knowledge of the future is uncertain. Therefore a plan is an attempt to extrapolate a line of practical reasoning into the future, based on what is known of the past, and what can be expected to plausibly happen in the future. A plan considers different possible or plausible outcomes (Wilensky, 1983).

Many actions require agents to collaborate, and to act together. In some cases, a group, like a corporation or a football team, act like what is, in effect, a single agent made up of many individuals acting as a unit (Castelfranchi, 1995). If the company declared bankruptcy, our way of describing this action implies or suggests that the company acted as the agent of the action taken. Or if the team won the game, we may ascribe the action of winning to the team as a whole, and not to any single individual. In any event, when agents act together collaboratively, as in a team, each participating agent will have a role, or part in the execution of the action. To take part in such a collaborative effort, each agent must have expectations about what other agents are assumed to do, and to be able to act on a basis of being able to guess how that other agent is likely to act in standard kinds of situations. This ability could be called anticipation.

A final capability an agent must have is that of deceptive action. An agent must be able to take advantage of the expectations of another agent, by not only anticipating how that other agent can reasonably be expected act. She must also be able to take advantage of that capability, and to deceive the other agent. Deceptive action involves the ability to take advantage of another agent's anticipation that you (as agent) will act in a certain way, in order to make that other agent think you are going act in that way, and then act in another way that the other agent does not anticipate. Deception is a complex kind of action, however, and further discussion of it follows below.

For a participant in a dialog to have the characteristics of an agent discussed above, the participant would have to have the following ten capabilities.

1. The capability to carry out actions.
2. The capability to have goals.
3. The capability to have information about its circumstances.
4. The capability to formulate and change goals.
5. The possession of long-term qualities like honesty and integrity.
6. The capability to have feedback on immediate consequences of its actions.
7. The capability to have reasonable expectations on the way things would normally be expected to go in a familiar type of situation.
8. The capability of drawing plausible inferences by implicature.
9. The capability of anticipation of moves of a dialog partner.
10. The capability for deceiving a partner in dialog by making moves that are misleading to the partner.

Can all of these capabilities be modeled by extending the current dialectical systems to have certain formal features? How could it be done? Basically what is required is to have a participant that is capable of carrying our practical reasoning by performing actions based on its goals. Specially marked commitments would be designated as representing an agent's goals. Some method needs to be given for determining whether the agent's having carried out a particular action counts as living up to her commitment. Such matters have already been widely studied in the literature on action theory, and within a formal theory of imperatives, the notion of "living up to a commitment" has been studied by Hamblin (1987). Such notions fit comfortably into the framework of the deliberation type of dialog. But how could they be modelled in the persuasion type of dialog? That seems to be the central problem, as far as the use of formal systems of dialog for the study of fallacies is concerned. How could we enrich the notion of a participant in the persuasion dialog so that it could have at least some of the key properties of an agent?

Another extension required is that the participant in dialog must have the capability of not only receiving information, but also of processing that information and acting on its consequences. This capability represents a whole group of special attributes. Part of accommodating this group of attributes could be achieved by expanding the commitment set to include a subset of propositions representing information that has come in – empirical information of a kind that the participant thinks to be accurate based on what she has observed. This subset could be called an information set. Information sets would be vital in information-seeking dialog, but would also play important roles in other types of dialog like deliberation and persuasion dialog.

Capabilities 6 and 7 relate to the plausible reasoning of a participant in dialog. The participant must not only be able to draw elementary inferences using deductive and inductive logic. The participant must also be able to make plausible conjectures (hypotheses), to draw plausible conclusions from these hypotheses, and to be aware of how the other party in a dialog will also draw such conclusions. Adding in these capabilities is not too hard. What is needed is a framework for plausible reasoning.

Finally, the participant needs to be modeled as possessing the capability for deception. This capability requires several prior skills, represented by the previous capabilities in the list of eight above. The deceiver needs to anticipate how the other party will likely respond to a certain sort of move, or sequence of moves in a dialog exchange. Then the deceiver needs to exploit this anticipated reaction by following it up with other moves that contribute to the deception. The deceiver needs to be able to anticipate how the other party will likely react to a certain move or development in a typical or standard kind of situation that both are familiar with. The deceiver needs to judge how the other party will draw plausible conclusions from premises that he thinks are true. Indeed, as we go over the list of capa-

bilities above, it becomes apparent that making deceptive moves of the kind associated with fallacies requires all ten prior capabilities.

7   Expectations and plausible inference

A lot of practical reasoning is based on the agent's expectations of how things normally go in a familiar kind of situation. For example, in the wood cutting action case, the agent expects that when he pulls the cord, the saw will start. It may not start, and then the agent may seek an explanation. If he smells gas, he may draw the conclusion that the motor is flooded. What to do? He may decide to let the saw sit for a while, hoping that the excess gas will drain out of the carburetor. Much of the reasoning in such cases is based on what is called inference to the best explanation.

With respect to this kind of familiar practical reasoning, there are many questions about whether it is based on probability, in some sense, or some other kind of inference that is different from probability in the statistical sense. What may be observed from any case is that such reasoning is based on what may generally be expected to happen, subject to exceptions, in a familiar kind of situation. This whole area of study is associated with default inferences and abductive reasoning. Some would say, however, that its origins are ancient. The Greek philosophers, like Plato and Aristotle, were familiar with it as a distinctive kind of reasoning called plausible reasoning – traditionally but misleadingly translated as "probability" in the rendering of Greek and Latin writings into English (Gagarin, 1994). It represents a kind of reasoning on the basis of what seems to be true in a familiar kind of situation, based on how one would normally expect other agents to act in that situation. The classic case is a lawsuit in which a smaller man, accused of assaulting a larger man, asked the jury whether it is plausible to assume that he, the smaller and weaker man, would attack this other larger and stronger man. The jury finds such a hypothesis not very plausible. Why? Well, the answer is (plausibly) that they can use empathy and put themselves into the situation of the smaller man in the given situation. Would they attack the larger man, if they were the smaller man in that situation? No, probably not, unless there were other overriding considerations in the case. Attacking the larger man would be imprudent, and probably even futile. Of course, new information in the case might defeat this plausible inference. For example, if the smaller man were really enraged, or if he was known to be an expert in boxing, the plausibility of the inference would be defeated, or at least reduced.

Another example is the commercial in which a highway patrol officer is talking to a husband whose wife has gone missing. The officer shines his flashlight into their car found abandoned at the roadside, and finds a wrapper on the front seat. The husband says, "What's that?" The officer replies, "Looks like the remains of an

individually wrapped Pepsid-AC Chewable." The two men then walk over to a roadside diner that specializes in hot, spicy food, and there they find the missing woman, eating some spicy food. Initially, when the officer indicated the diner as a place to look, the husband told him that his wife suffers from heartburn and cannot eat spicy food. This information went against the hypotheses that the women had gone to the diner. But when the officer found the "remains of an individually wrapped Pepsid-AC Chewable", a chain of plausible reasoning was suggested as a better explanation of what likely happened. Once the woman had taken the individually wrapped Pepsid-AC Chewable, she was then presumably able to eat spicy food. Therefore, since it was possible for her to eat spicy food, and since the diner is the place you go for spicy food, it is plausible to assume that she may have gone to the diner. Once the pair of men enter the diner, their hypothesis is confirmed. There is the woman, eating some spicy food.

Formal systems for plausible reasoning have been proposed by Rescher (1976). In these systems, plausible inferences are determined by how things can be expected to go in a normal or typical situation familiar to the reasoner. Rescher shows how the system of calculation for plausible inference should be seen as different from the model of rational inference presupposed by the probability calculus. Rescher's system of plausible reasoning is applicable to the inferences drawn in the Pepsid-AC case. The two men draw their conclusion to take action by looking in the diner by making up a hypothesis based on their inference to the best, or most plausible explanation of what most likely happened. The given evidence of the wrapper on the car seat suggests a chain of plausible reasoning. The unstated presumption exploited by the commercial is that both men know that the woman knows that taking Pepsid-AC makes it possible to eat spicy food without unpleasant consequences. Another unstated assumption is that when both men see the diner, they know that it is a place where you can get food. On this basis, they take the action of walking to the diner. When they get there and look in, their hypothesis is confirmed. The plausible conclusion they drew turned out to be right.

Plausible reasoning is highly characteristic of not only how agents reason, but how they engage in collaborative argumentation in dialogs with other agents. Plausible reasoning by an agent is based on the agent's ability to use empathy to understand the reasoning of another agent, based on observing how that agent acted in a situation that is familiar, and in which conclusions may be drawn on the basis of what would be normally expected to happen in that kind of situation. Such a familiar situation is called a script in the literature on artificial intelligence (Schank and Abelson, 1977). As an example of a script, they use the case of someone going to a restaurant, sitting down, picking up the menu, ordering some food, eating the food, paying for it, and then leaving (pp. 42–46). All of us understand this familiar sequence of actions and know how it typically goes. Such an example

of a script shows that what is deeply characteristic of an agent, in the kind of case where two or more agents are engaged in reasoning together in a dialog. It is the capability of the agent to have expectations about the way that things can be expected to go in a normal or familiar type of situation. An agent can anticipate how another agent is likely to react, or should normally react, in a familiar kind of situation. The agent can then draw presumptive conclusions on this basis, using plausible reasoning based on shared knowledge of a familiar kind of situation. The script is a contextual notion. We can fill in the gaps in a sequence of actions in a script, because we are familiar with this type of action, and how it fits in with conventional daily activities we have been through before. In a case of argumentation, the dialog is also a contextual notion that enables us to fill in missing steps in a sequence of argumentation using plausible reasoning.

8   Plans, strategies and chaining forward

The key feature in the formal structure of a deliberation type of dialog is that each agent must have a plan. A plan is a special part of an agent's commitment set in which she sets out her plausible hypotheses about what is likely to happen in the future relative to a specific situation that is developing. But in a persuasion dialog, instead of a plan, a participant has a strategy for persuading the other party. But a strategy of persuasion can be seen as a species of plan. In a persuasion dialog, an agent tries to anticipate how her dialog partner will plausibly react to future moves in the dialog, insofar as her planned sequence of argumentation, and the other party's responses to it, can be extrapolated forwards. The plan or strategy is used to make up argumentation sequences of attack and defence against a specific opponent, to the extent that the opponent's goals and other commitments are known or can be anticipated. There are two techniques used in the formation of such a tactical plan of dialog argumentation. One is the chaining forward of inferences. The other is the guessing or anticipating of the future sequence of moves in the dialog. To make a strategy, an agent in a dialog tries to extrapolate the line of argumentation forward, to the extent that this can plausibly be forecast.

Such a tactical action strategy in a dialog can be modeled as a sequence of argumentation that is like any typical dialog sequence of argumentation, except that the sequence has not really been played out in actual dialog. It is a sequence that exists only in the agent, as a hypothetical dialog that might be played out in the future. It is a sequence that might actually occur in the future playing out of the actual dialog, but has not yet actually occurred in this particular dialog. Or then again it may not. The agent who has formulated the strategy hopes that it will work out, and that the ultimate move in the sequence will be the fulfillment of her goal

in the dialog. What must a participant do in a persuasion dialog to formulate a strategy to get the other party to come to accept her thesis? What she must do is to take the statements she knows to be commitments of the other party, and then, using plausible reasoning and implicature from this basis of premises, guess what other statements the other party would also likely be committed to. Then she must build up a chain of argumentation, using only these estimated commitments as premises, that eventuates in the thesis to be proved.

In certain respects, a strategy in a persuasion dialog would be very much like a strategy for action of the kind that is used in a deliberation type of dialog. There are formal theories of action in which the notion of an action strategy has an important place. Segerberg (1984) has developed a formal theory of action in which actions are analyzed as made of subactions arrayed in a sequence of the kinds described by Goldman (1970). An example of running a routine would be the action of cooking a meal based on a recipe (Segerberg, 1985, pp. 185–187). The recipe could be thought of as a kind of action strategy, and the actions of cooking the meal are the running of many routines determined by the recipe. Hamblin (1987) has constructed a formal theory of action strategies as part of his logic of imperatives. Hamblin (p. 155) defined a strategy formally as " an allocation of deeds, one for each time, and for each appropriate context at that time." Hamblin defines a deed (p. 140) as an elementary independent action attributable to a particular agent. As Hamblin notes (p. 157), a strategy does not specify every deed in minute detail. As he puts it (p. 157), "some deed-choices are left open." A partial allocation of deeds is called a partial strategy (Hamblin, 1987, p. 157). We can think of a partial strategy (Walton and Krabbe, 1995, p. 17) as a set from which an agent should select one deed at a time in order to follow an action strategy, given what the agent thinks was the history of the world up to the time of carrying out the action. In other words, a partial strategy specifies a sequence of actions with gaps in it. The gaps can later on be filled in as events actually unfold in the given situation in which the action is being carried out. A formal precis of Hamblin's action-state semantics has been given in (Walton and Krabbe, 1995, pp. 189–195). The notion of a partial strategy is fundamental to understanding the notion of commitment, as shown in Walton and Krabbe (1995, pp. 15–21). A person may rightly be said to be living up to a commitment as long as her actions indicate a partial strategy and a sequence of actions that lead towards the fulfillment of that commitment and do not lead away from it. But the notions of leading towards and leading away have to be evaluated in specific cases in relation to how the partial strategy, as reconstructed in the case, fits in with the known actions in the case.

So how does this formal notion of an agent's strategy for action fit with the notion of argumentation strategy required to analyze and evaluate cases like the critical discussion on euthanasia? This case is not one of a joint deliberation dialog in

which the two participants are trying to decide on a prudent course of action in a given situation. Actions and consequences are involved, as indicated in section 1 above. But the case is a persuasion dialog in which each party has a thesis, or conclusion to be proved, and each side tries to use the strongest and most persuasive arguments to prove that thesis to be true. The strategies involved in this dialog are not strategies of action, but strategies of reasoned persuasion using arguments. Basically, in a persuasion dialog, two arguers ask questions and put arguments to each other, where the aim is to get the other party to become committed to propositions that can be used to prove one's own thesis by a connected chain of argumentation. Each individual inference in a chain of inferences is supposed to be valid, or structurally correct, according to the forms of argument appropriate for use in this type of dialog. In a persuasion dialog, various kinds of moves are allowed, including the asking of questions, the answering of these questions, and the putting forward of arguments. When it comes to the putting forward of arguments, there are four basic requirements that determine what is an argument that has been used successfully and appropriately by a proponent in the dialog, to prove a conclusion.

(*R1*) The respondent accepts the premises as commitments.

(*R2*) Each inference in the chain of argument is structurally correct.

(*R3*) The chain of argumentation must have the proponent's thesis as its (ultimate) conclusion.

(*R4*) Arguments meeting (*R1*), (*R2*) and (*R3*) are the only means that count as fulfilling the proponent's goal in the dialog.

The practice of using purely hypothetical arguments may seem to violate (*R1*), but really it does not, because in the end, for an argument to count as rationally persuasive to a respondent, it must be based on premises that he has come to accept (as commitments) in the dialog. (*R2*) and (*R3*) depend on the possibility of chaining together a sequence of inferences in argumentation. Consider the following example of a *modus ponens* type of inference based on a conditional proposition and a simple proposition.

(*I1*) If *S* then *T*

    S

    ————————————

    T

This inference can be "chained together" with another one–say, for example (*I2*).

(*I2*)   If *T* then *U*

T

——————————

U

The two inferences (*I1*) and (*I2*) are chained together because the conclusion of (*I1*), *T*, is utilized as a premise in (*I2*). Chaining together of subarguments into a longer chain of argumentation is also a familiar phenomenon in argument diagramming. A chain of argumentation can be modeled as a pathway of reasoning in an argument diagram using the method of argument diagramming in (Walton, 1996, chapter 6). In this method, a *line of reasoning* is an alternating sequence of propositions and inference-steps where each step goes from one proposition to the next. A *pathway of reasoning* is a line of reasoning in which all the propositions are distinct (p. 189). A pathway of reasoning, in other words, is a line of reasoning in which there is no circular line of reasoning. The pathway of reasoning can be used to anticipate where a line of argumentation seems to be going, in a given case toward some ultimate conclusion, based on the line of evidence indicated in the pathway so far in that case.

In addition, an agent needs to be seen as an entity that has reasonable expectations about the way things can normally be expected to go in a familiar type of situation, and can draw conclusions, using plausible reasoning, either when these expectations are confirmed or when they are violated. Such expectations can also be seen as statements and rules (conditionals) that are included in an agent's set of commitments. They are not statements the agent is fully committed to, but only statements she is committed to as being generally true, or plausibly acceptable, subject to retraction if contrary information comes in. These kinds of commitments are defeasible. They are tentatively accepted, subject to defeat in the face of contrary information, or if it turns out the given case is an exception to the rule. An agent will have expectations about what is likely to occur, but these expectations are only based on standard patterns of what is typical. They may turn out to be false. A strategy in a persuasion dialog needs to be based on not just chaining forward of deductively valid inferences, but on chaining forward of plausible inferences.

All these considerations show how argumentation in a persuasion dialog can be seen as, in certain key respects, similar to the kind of goal-directed reasoning typically found in agent deliberation. The participant in the persuasion dialog is seen as making moves that anticipate the other party's moves by devising a strategy, like a plan of action. The participant chains argumentation forward, devising what Hamblin calls a partial strategy. The projected sequence of argumentation is incomplete, and the gaps are not filled in until the strategist actually comes to grips

with the opponent when the strategy is acted out by actually putting argumentation forward. To see a strategy of argumentation as a kind of plan, what is required is to view the arguer as an agent who has goals, that carries out speech acts, and that has information about its circumstances, including feedback on the immediate consequences of its actions. Modeling a participant in a persuasion dialog as an entity having goals is not a big problem. Goals can be represented as statements (general or specific) that are included in the agent's commitment set. Another factor to be modeled is the transmission of information. What about the agent's possession of information about its circumstances, and use of that information in making a plan? This feature is central to the information-seeking type of dialog, but is more peripheral in persuasion dialog. In a persuasion dialog, this feature can be added to the commitment set, if we see the agents in a dialog as agreeing on certain "facts" as statements they both accept as true, and agree not to dispute.

9   Strategies in formal dialectic

To analyze and evaluate the argumentation in the case of the critical discussion on euthanasia (section 1, above), it is useful to see Wilma and Bob not just as passive followers (or violators) of the rules of the critical discussion type of dialog. It is much more useful to see them as having certain key properties of agents. They can not only make moves (speech acts) in the dialog, but they can anticipate the moves of the other party. The other party can then react to those moves by using strategies (plans) that had been devised earlier for just this sort of contingency. The participants can even use tricky deceptive tactics specifically designed in advance to appear attractive, based on what is known about the commitments of the other party. Such a participant can not only act, but can act proactively by anticipating the plausible or expected moves of the other party. At the same time, an agent can use feedback to modify her plan of attack as the other party reacts to it, move by move. When agents use a combination of feedback and planning in this way, they are engaging in practical reasoning, and implementing a strategy or plan. The plan can be partly represented as a goal, or set of goals. But as attempts to carry out the goal are made, ways and means of carrying out the plan are modified as new information comes in, and as the consequences of the previous actions to carry out the goal are observed. This typical pattern of practical reasoning represents what happens when one party in a critical discussion presents a subtle argument, like a slippery slope argument, and anticipates how the other party is likely to react to it.

Wilma's slippery slope arguments, as Bob observed, makes a big issue of the unintended bad consequences that could flow from implementing a euthanasia policy. Wilma's argument assumes that euthanasia is dangerous because it is policy

that has potentially bad consequences, and has not been tested on a wide scale. As shown in section 1, Wilma's line of argumentation is based on argument from negative consequences. Her argument postulates two premises. One is the conditional that if physician-assisted suicide is allowed as a practice, bad consequences will follow. The other is that we should not allow these bad consequences to occur. The conclusion that follows is that physician-assisted suicide should not be allowed as a practice. This argument, because it has the form of argument from negative consequences, already shows how the argumentation in the critical discussion on euthanasia case can be modeled. Bob and Wilma need to be seen as agents engaged in using practical reasoning in their arguments with each other. But modeling both participants as agents using argumentation strategies can be seen to be even more useful when it comes to analyzing the case. Bob's strategy was to hold back his rebuttal on the basis of the Netherlands case, and wait until the right moment in the dialog arrived. Bob also strategically anticipated how Wilma might plausibly react to that counter-argument, and constructed some additional arguments that could be used as plausible rebuttals to her counter-arguments.

Strategy certainly is based on chaining forward of a line of argumentation, but that is not all there is to it. It also involves knowledge of how certain commonly used forms of argument (argumentation schemes) work. Each argumentation scheme has a matching set of critical questions. Asking the right critical question at the right point in a dialog can open up opportunities. These opportunities can be exploited by holding a powerful argument in reserve for just the right moment. In rhetoric, the use of such strategies is associated with the expression "the opportune moment" (*kairos*). To exploit the opportune moment, a strategy has to be flexible, because you can never know in advance what moves in the dialog your opponent will actually make. But good strategy requires tailoring your own moves to those of your opponent.

Bob's strategy in the euthanasia case involves two kinds of tactics. One is wait until Wilma brings out the slippery slope argument and then ask critical questions that will make her fill in some of the gaps in the argument. Then once the right gap is filled by Wilma's making a particular claim, Bob can use his Netherlands counter-example. Suppose the dialog takes the following form.

Wilma: Once euthanasia gets going, there will be no stopping the killings, and ultimately we will get Nazi death camps.

Bob: Hang on a minute. Could not physician-assisted suicide be restricted to those who are terminally ill, and who voluntarily ask for it, in consultation with a physician they know quite well. How do we get from there to Nazi death camps?

Wilma: Well, once any form of killing is accepted as social policy, there is no resisting the slippery slope. There is such a loss of respect for life that sets in. One thing just leads to the next.

Bob: But how do you know?

Wilma: Well that's exactly what happened in Nazi Germany. It all started with their euthanasia program.

Bob: Yes, but that's different. It was a dictatorship run by a murderous fanatic who had all kinds of crazy racial theories about survival of the fittest, and so forth. That is not the situation to the current western democratic countries.

Wilma: Yes, but once this killing gets started in any form, even in a democracy, the policy of euthanasia can be used to kill off persons who are judged to be socially undesirable, even if they do not agree to euthanasia voluntarily.

This is exactly the point where Bob can use his Netherlands counter-example. He can then argue that there is a western democracy that has now had a physician assisted suicide policy in effect for several years, and that the program has been successfully restricted to persons who are terminally ill and who voluntarily ask for it.

To grasp Wilma's line of argumentation, and how Bob reacts to it by posing critical questions and counter-arguments, we have to see that two argumentation schemes are centrally involved. The first is the scheme for the slippery slope argument.

Scheme for the slippery slope argument (Walton, 2006, p. 107)

FIRST STEP PREMISE: A_0 is up for consideration as a proposal that seems initially like something that should be brought about.

RECURSIVE PREMISE: Bringing up A_0 would plausibly lead (in the given circumstances, as far as we know) to A_1, which would in turn plausibly lead to A_2, and so forth, through the sequence $A_2,... A_n$.

BAD OUTCOME PREMISE: A_n is a horrible (disastrous, bad) outcome.

CONCLUSION: A_0 should not be brought about.

Critical questions for the slippery slope argument (Walton, 2006, p. 110)

CQ_1:   What intervening propositions in the sequence linking up A_0 with A_n are actually given?

CQ₂: What other steps are required to fill in the sequence of events, to make it plausible?

CQ₃: What are the weakest links in the sequence, where specific critical questions should be asked, on whether one event will really lead to another?

Wilma uses the slippery slope argument when she says that once euthanasia gets going, there will be no stopping the killings. Her argument adds the bad outcome premise that ultimately the chain of reasoning will lead to the Nazi death camps situation. The conclusion is that physician-assisted suicide should not be brought about as a general policy. But the recursive premise in her argument is very weak in this instance. Bob raises this question by asking how we get from the present situation to the Nazi death camps situation. She has postulated quite a long sequence of argumentation, and failed to fill in the gaps needed to make it plausible. Bob proposes a counter-argument by presenting a policy that might stop this chain of argumentation from leading to its ultimate conclusion. He asks whether physician-assisted suicide could be restricted to those who are terminally ill and who voluntarily ask for it. To respond to this criticism, Wilma inserts another argument as a link into the chain of reasoning, citing loss of respect for life as a factor that generates the sequence of events in the chain of reasoning. Bob questions this attempt to reinforce the argument by simply asking for evidence that one thing leads to the next. This part of the dialog could be called the first stage.

The second stage involves another argumentation scheme. Wilma brings in an argument from analogy claiming that the sequence of reasoning at issue in the present situation is comparable to the historical case of the euthanasia program of Nazi Germany.

Scheme for argument from analogy (Walton, 2006, pp. 96–97; Walton 1989, p. 256)

SIMILARITY PREMISE: Generally, case *C1* is similar to case *C2*.

BASE PREMISE: *A* is true (false) in case *C1*.

CONCLUSION: *A* is true (false) in case *C2*.

Critical questions for argument from analogy

CQ₁: Are there differences between *C1* and *C2* that would tend to undermine the force of the similarity cited?

CQ₂: Is *A* true (false) in *C1*?

CQ₃: Is there some other case *C3* that is also similar to *C1*, but in which *A* is false (true)?

At his next move, Bob argues that the argument from analogy fails, because the one situation is not similar enough to the other to warrant the conclusion drawn. The focus of this attack is the first critical question above. At her last move, Wilma replies that the slippery slope argument still works, because even in a democracy the policy of euthanasia can be used to kill persons who are judged to be socially undesirable. She does not cite a comparable case, but the way is left open for her to do so. The strategic avenue left open here is for her to attack Bob's attempted refutation of her argument from analogy by using a counter-analogy.

What kind of strategy has Bob used in this kind of case? Several elements are involved. One is that Bob has a good idea of where Wilma's line of argumentation is going, he has some grasp of how slippery slope arguments generally work. He knows that her aim is to justify her thesis that euthanasia is a bad policy using argumentation from negative consequences, but in that special form we know as the slippery slope argument. He anticipates that she will use the slippery slope argument to support her ultimate conclusion that euthanasia should not be set into place as a general policy. Assuming that Bob has some grasp of how slippery slope arguments work, he may know that the slippery slope argument is hard to support adequately by good evidence, and that it is open to certain kinds of critical questions and counter-arguments. He does not know how Wilma may reply to such critical questions, but he anticipates that when she is forced to reply to them, that may give him the right opening to present his Netherlands counter-example most effectively. How these elements of strategy are combined in the case of the critical discussion on euthanasia case is illustrated by figure 5.1, where a supportive argument is represented by a line with a single arrowhead, the direction of the arrow indicating the direction of the inference from the premise or premises to the conclusion. A refutation taking the form of an attack on a prior argument or statement is represented by a line with a double arrowhead. A refutation can be a critical question that raises doubts about a prior argument or it can be a counter-argument.

In addition to containing a proposition or question, each box contains an owner of that speech act, namely Wilma or Bob. The only exception is the box containing the label of the slippery slope argument. It represents the argumentation scheme that is the basis of Wilma's argument. Bob's first question raises doubts about Wilma's previous argument, and is therefore connected to it by a double arrow. Bob's second question, further to the right in figure 5.1, is a rhetorical question that backs up his first question by supplying a guideline that would refute Wilma's slippery slope argument. Wilma's reply to this criticism is presented in the two boxes just below it. Wilma argues that once a form of killing is accepted as a social policy, there is no resisting the slippery slope. This assertion is joined to Bob's prior move above it by a double arrow, indicating it is an attempted refutation of Bob's argument posed by his rhetorical question. Backing up Wilma's asser-

tion is the box below it citing loss of respect for life as a factor. It is joined to the box above it by a single arrow, indicating support. At the same time as it attacks Bob's counter-argument, Wilma's argument also supports her original slippery slope argument. Thus an arrow with a single arrowhead joining her argument to her original slippery slope argument is inserted in figure 5.1.

WILMA: Ultimately we will get Nazi death camps.

Slippery Slope Argument

BOB: How do we get from there to Nazi death camps?

BOB: Could not physician-assisted suicide be restricted to those who are terminally ill, and who voluntarily ask for it?

WILMA: Once euthanasia gets going, there will be no stopping the killings.

WILMA: Once any form of killing is accepted as social policy, there is no resisting the slippery slope.

WILMA: There is such a loss of respect for life that sets in.

Figure 5.1 Elements of strategy in the critical discussion on euthanasia case

A similar diagram could be drawn representing the argumentation in the second part of the dialog where the argumentation scheme involved is that for argument from analogy. There Wilma cites the case of Nazi Germany as a similar case to back up her argument that once killing is accepted as a social policy, there is no resisting the slippery slope. Bob raises doubts about this argument by citing differences between the two cases. There is no need to draw a second diagram to show how the second stage of the dialog fits in with the argumentation in the first because it is easy to see how that can be done once the first stage of the dialog has been represented in figure 5.1.

What is shown by this analysis of the argumentation in the euthanasia dialog is that even though an arguer never knows in advance exactly what moves his opponent will make, nevertheless because the argument she is likely to use has a standard form, as an argumentation scheme, strategic planning is possible to some extent. Each argumentation scheme has critical questions that are appropriate for it. Asking these critical questions at the right points can create openings that can

be filled by arguments held in reserve. So part of the strategy is to resist the temptation to use a good argument too early in the dialog. Instead of tipping off the other party too early, when she can have time to think up rebuttals, it is better strategy to keep the argument in reserve, for use when it will be most effective, and most surprising to the opponent.

Those studying invention strategies for debating, in the field of rhetoric, are already quite familiar with preparing to argue cases in just this way, in competitive debates. Much more detailed cases of debating could be presented that use strategies similar to the relatively simple ones illustrated in the euthanasia case. In logic, in the study of fallacies, we are not nearly so familiar with using longer cases of this sort. In textbook accounts of informal fallacies, very brief examples are typically used, in which much of the case, including detailed accounts of strategies and dialog interactions of the participants, are left out. This tradition needs to be changed, for a much fuller analysis of fallacies can be given if more of the dialog context of a case, including the strategies of the participants, can be included in examples.

In the case of the critical discussion on euthanasia, seeing how the two strategies of Bob and Wilma interact in the dialog is quite revealing. It shows not only what the weak points are in a slippery slope argument of the kind used by Wilma, and what the requirements are for fulfilling a proper burden of proof for this kind of argument. It shows also how the argument can be used as a sophistical tactic type of fallacy. It shows how Wilma as a planning agent can anticipate Bob's plausible and expected moves by seeing him as an agent, and by reacting in advance to his plausible moves. The case shows how typically, both participants in a critical discussion will engage in this kind of anticipatory planning, making feasible plans that can be modified and improved by being tailored to the specific responses of the other party as the strategy is played out in the actual dialog. The case study needs to display both participants in the dialog as agents who have plans, and who can be seen as carrying out these plans as they interact with each other in their speech acts in the dialog.

How then do we need to enrich the current models of formal dialectic to accommodate this new notion of a participant as an agent who is capable of making up a strategy in advance and then using it in a way that takes feedback into account as the strategy is implemented? The answer is that the participant needs to be seen as having not only a commitment set, but as also having the capability of devising an argumentation strategy. A strategy contains hypothetical sequences of argumentation moves by two parties in projected fragments of dialog. The strategy is made up by one party's chaining forward a sequence of inferences that are potentially useful for an aim in a dialog. Each such sequence is called a strategy of a specific participant. The strategy needs to be seen as separate from the actual sequence of the real dialog, as carried out in the actual moves of the two participants.

A strategy does not necessarily represent any actual dialog. It only represents a hypothetical chaining forward of a sequence of argumentation that anticipates how the other participant may or may not react if that argument were to be put forward in later dialog moves. According to Hamblin (1970), a commitment set can be seen as a list of statements, written on a blackboard, for example, or recorded on a tape. Each participant in a dialog has such a set, and new statements can be inserted or old statements erased as the dialog proceeds. Presumably, the commitment set of each participant is on view to the other participant. Strategies, however, cannot be directly viewed by the other participant, and can only be surmised on a basis of plausible guesswork.

It is also useful to have other components in a dialog, as well as commitment sets. There can also be information sets of statements, representing incoming information from some source agreed upon as reliable by the participants. For example, in the critical discussion on euthanasia, both participants might agree to accept any facts about euthanasia that can be verified by some source, like an encyclopedia, or a respected scientific study. For example, Bob may have collected his facts about physician-assisted suicide by reading a recent scholarly article by a group of physicians and scientific experts who had collected data on the situation over several years. He and Wilma might have agreed to accept as factual any findings reported by this group. Or at least, Bob may reasonably expect that Wilma would not very likely be inclined to dispute any statements vouched for by such a source. Like the strategy sets, information sets may only be visible to the one party in a dialog and not to the other. When such information is brought forward and deployed as premises used in argumentation by the one party, then of course it will become known to the other party. In this respect, information sets are like strategies. The difference is that an information set can be any sets of presumed facts that can be vouched for by an appropriate source, while strategies are connected sequences of possible future moves in a dialog.

A strategy then is based on the following ten components. Each of these components has already been illustrated in the case of the discussion on euthanasia, or in other cases mentioned.

1. Information about the relevant facts of the case the participant has collected.
2. Information the participant has about the other party's commitments.
3. Awareness of argumentation schemes (forms of argument) that can be involved in argumentation used by either side.
4. The capability to chain forward to anticipate where the sequence of argumentation is likely to lead.
5. The capability to fit the commitments of the other party into such a sequence of future argumentation as premises.

6. The capability to anticipate key arguments that will be used at some future point by the other party.
7. Based on an argumentation scheme, the ability to ask appropriate critical questions in reply to a key argument, once that argument has been put forward by the other side.
8. The capability to present counter-arguments, especially arguments based on information collected about the particulars of a case.
9. The capability to hold back using a strategy until the right point in a sequence of dialog is reached.
10. The capability to grasp the ultimate thesis aimed at as the end point of the argumentation in the dialog, and to judge whether argumentation is moving towards or away from that thesis.

The strategy is made up of these capabilities and kinds of knowledge, as applied to a given case. An actual strategy adopted in a given case could be seen as the results of collecting this information and applying these capabilities, in relation to the dialog in the given case. For example, in the critical discussion on euthanasia case, both Bob and Wilma may have in mind, or may even have drawn up on paper, their conjectures about where the argument will likely lead, and how to counter or reply to moves they anticipate the other side may make. A set of notes of this kind could function as a very clear representation of a strategy. The notes could contain items based on some or all of the ten components. A brief strategy would have only some components. A more complete strategy would have elements representing all ten components. Each component could be elaborated in more or less detail, depending on the specifics of the case.

A strategy could be visualized as a small blackboard or piece of paper possessed by each participant, containing a connected sequence of dialog moves. Each connected sequence of moves representing a strategy is especially tailored to not only a type of dialog, but to a specific kind of juncture or situation that might occur in that type of dialog. The statements in a strategy do not become actual commitments of a participant unless they are actually put forward as moves in the dialog. The strategy represents special kinds of argument moves that are hypothetical in nature. The strategy may never actually have to be put into action, so to speak, and so the arguer may never actually go on record as committing herself to the statements contained in the plan. However, she would be ready to use the strategy, should the right occasion arise. And so she can be committed to the plan, and to the arguments and moves required to carry it out.

Once strategies are recognized as distinctive entities in systems of formal dialectic, many different types of argumentation strategies can be studied and classified. To indicate how such a project could be fertile ground for further work, two

types of strategies could be cited. The strategy of diversion has already been associated (in section 2 above) with fallacies of irrelevance, like the red herring fallacy. Not all argumentation strategies that are interesting from a logical point of view are associated with traditional informal fallacies, however. The use of persuasive definitions, and the tactical use of emotive language generally, can nicely be explained using deceptive concealment as an argumentation strategy. The analysis of persuasive definitions of Stevenson (1944) showed how the use of a persuasive definition can exploit the existing emotive meaning of a term when that term is redefined. For even though the term has been given a new definition, supposedly representing the "true nature" of the thing defined, the old existing meaning lingers on in common usage. And this old existing meaning can have a positive or negative emotive spin that influences an audience towards acceptance or rejection.

10   Qualities of character for agents in formal dialectic

We have now reached the point where the participants in a persuasion dialog can be seen as having many characteristics comparable to those of an agent in a deliberation dialog. But how far can we go with this view of the participant in the persuasion dialog? One further step would be to see the participant in persuasion dialog as having qualities of character like honesty, integrity and sincerity. But some might think that this step is going too far. Naturally, some would question why there is a need to model qualities of character at all, in order to evaluate argumentation of the kind associated with informal fallacies. One supreme advantage of taking this controversial step is that arguments depending on the credibility of a source could be evaluated.

Of particular interest with respect to the study of fallacies, *ad hominem* arguments and *ad verecundiam* arguments could be modeled and evaluated much more realistically. By *ad hominem* argument is meant the use of personal attack on the character of an arguer in order to run down or devaluate his argument (Walton, 1998). By *ad verecundiam* argument is meant the appeal to expert opinion to support (or undermine) an argument (Walton, 1997). Both types of arguments depend on the credibility of an arguer or source cited, and this in turn can depend on the perceived character of the arguer. For example, the testimony of an expert witness may be attacked in court by arguing that the witness has a bad character for veracity. Generally, the testimony of any witness in court can be attacked on the grounds of the character for veracity of the witness (FRE, 1987). To be able to fully analyze and evaluate *ad hominem* arguments – if this type of argument is seen as a personal attack on an arguer's character – the notion of character must be somehow represented. This shows not only that *ad hominem* arguments and ap-

peal to expert opinion arguments are sometimes quite reasonable in legal argu-
mentation. It also shows how the evaluation of such arguments properly depends
on the character of the arguer. Even when such arguments are judged to be falla-
cious, surely the evaluation of the argument needs to depend on considerations of
the arguer's qualities of character.

The argumentation for the basic type of *ad hominem* is called the generic *ad
hominem* in (Walton, 2006, p. 213).

Scheme for the generic type of *ad hominem* argument

> CHARACTER ATTACK PREMISE: *a* is a person of bad character.
>
> CONCLUSION: *a*'s argument α should not be accepted.

Critical questions for the generic type of *ad hominem* argument

> CQ$_1$:   How well supported by evidence is the allegation made in the charac-
> ter attack premise?
>
> CQ$_2$:   Is the issue of character relevant in the type of dialog in which the ar-
> gument was used?
>
> CQ$_3$:   Is the conclusion of the argument that α should be (absolutely) reject-
> ed, even if other evidence to support α has been presented, or is the
> conclusion merely (the relative claim) that α should be assigned a re-
> duced weight of credibility as a supporter of α, relative to the total
> body of evidence available?

What kind of strategy is involved in using *ad hominem* arguments successfully?
First, such strategy may involve collecting information about your opponent's char-
acter. But the most crucial part of the strategy is that of figuring out how to use this
information maximally at the right time in a dialog. An attack on an arguer's char-
acter at just the right point in a dialog can have devastating consequences, even
though the same attack, at some other point in the dialog, or if used in a different
context of dialog, might be weak, or may even backfire, making the attacker look
bad (Walton, 1998). When an *ad hominem* argument is successful, it is because it has
effectively undermined an opponent's credibility, just at a point in the dialog where
her argumentation depends crucially on that credibility. An arguer's credibility de-
pends on her *ethos*, that is, her perceived character or reputation (Brinton, 1986).
For example, if an arguer has a positive *ethos*, because she is an expert in the domain
of knowledge into which the argument falls, and because she has a personal reputa-
tion for honesty and integrity, that positive credibility will, quite appropriately, tend
to make her argument from expert opinion more plausible. But what if that expert

can be shown to have committed perjury, or lied under oath, in the past? A critic can use that information to cast doubt on her credibility, but the key is to use it just at the right moment, when attacking her credibility as a trustworthy expert will have the greatest impact in undermining the plausibility of her claims.

Timing is the key in using such *ad hominem* arguments. Even a weak and poorly substantiated *ad hominem* attack can have explosive impact in a dialog, tilting the whole balance against the side that was attacked. The side that was attacked needs to have defensive strategies. But having a counter-strategy in place can be problematic if the other party has taken tactical advantage of surprise. Matters of strategy in using *ad hominem* arguments are quite similar, in certain respects, to strategic uses of the slippery slope argument, as illustrated in the critical discussion of euthanasia case in chapter 5, section 9. On the one side, it is a matter of using the argument at just the right point in the dialog. On the other side, it is a matter of devising a strategy to counter that argument, but waiting until the right moment, depending on how the other side moves when they bring forward the argument. Both sides can use strategies, but both sides need to tailor their strategies to how the other side moves.

Evaluating *ad hominem* arguments as fallacious or not can be greatly assisted by having a grasp of various strategies, and how they are used. One can distinguish between cases where the problem is one of relevance, and cases where the *ad hominem* attack is relevant, but the problem is that the premise claiming bad character is weakly or insufficiently supported. One kind of critical question is to ask whether the *ad hominem* argument is relevant in the dialog. Another critical question is to ask whether the premise alleging bad character is supported by strong evidence. As in all cases of evaluating fallacies, strategies are linked to the critical questions that are appropriate for a given argumentation scheme. But the main thing to notice here is that *ad hominem* arguments can only be adequately understood, analyzed and evaluated if the arguer who is attacked is seen as having qualities of character that affect how we judge his credibility. But the best way of making sense of the notion of a participant in dialog having such qualities of character is to see that participant as an agent. The importance of seeing a participant in dialog as an agent particularly comes to the forefront when evaluating cases of the use of the circumstantial *ad hominem* argument, where a participant is attacked for not practicing what he preaches. Here what is crucial is the presumption that the participant has carried out certain actions in the past, and the claim is that his arguments at this point in the dialog are inconsistent with those past actions. The upshot, as in all *ad hominem* arguments, is an attack on the arguer's credibility.

Even over and above the project of analyzing informal fallacies, there are other reasons for thinking that seeing participants in dialogs as agents represents an interesting avenue of exploration. One reason is that qualities of character, like

honesty and sincerity, have been shown to be centrally important to conversa-tional postulates by Grice (1975). Another reason is the new technology of multi-agent systems (Wooldridge and Jennings, 1995) has shown how conversational policies that agents must use for collaborative actions depend on assumptions one agent makes about the character of another agent. Agents have the capability not only to act and react in specific situations. They can even have qualities of charac-ter that are stable over long periods (Wooldridge and Jennings, 1995). For exam-ple, an agent could be said to be honest if she generally appears to try to tell the truth, and if she has no record of lying over a protracted period of engaging in a conversation. Or an agent could be said to have integrity if his actions are consist-ent with his arguments, and what he advocates generally, over a long period of time. Honesty and integrity are relatively stable qualities of character (Brinton, 1986). These are the kinds of long-term character traits that agents have, or fail to have. The problem is how agents, and qualities of character that agents are thought to have, like honesty and integrity, should best be modeled in systems of formal dialectic. The present systems of formal dialectic have two participants – a propo-nent and a respondent. Both participants are holders of commitment sets. But how could qualities of character like honesty and integrity be incorporated? Could they somehow be inserted into the commitment sets? The proposal that seems best is to have special subsets of the commitment set that contain assertions about long-term stable qualities of character the participant is known to have. These commit-ments could then be modified as relevant information on such matters comes in through the progress of the dialog. It is possibly too soon to contemplate exactly how such qualities of character would be evaluated and utilized in different types of dialog. To model *ad hominem* and *ad verecundiam* cases, it is necessary to see each participant as having a given credibility value, and then to see this value being raised or lowered as the result of certain kinds of moves in the dialog. However, it is already clear from the literature on these fallacies (Walton, 1997; 1998) that some structure of this sort is what is required.

Dialectical shifts and embeddings

A subject that has been little studied in dialog theory is the dialectical shift, or change from one type of dialog to another during a sequence of argumentation. Dialectical shifts were studied in (Walton and Krabbe, 1995), but very little systematic work has been done on them since that time. Shifts are important for dialog theory because they are very common in argumentation, and because they are often associated with fallacies and other significant problems of argumentation (Walton and Krabbe, 1995; van Laar, 2003). One type of shift that can sometimes be highly problematic, as shown in section 1, is the burden of proof shift. In some shifts there is an interruption of the ground level dialog when the shift occurs, so that the advent of the second dialog is an easily visible break from the line of argumentation in the first dialog. In other cases, the dialog seems to flow smoothly along over the shift so that the second dialog fits nicely into the first. This second type of case, called a dialectical embedding (Walton and Krabbe, 1995, p. 102), is said to occur where there is a productive functional relationship between the two dialogs so that the argumentation in the second dialog enhances the quality of the argumentation in the first.

Little has been written on the problem of embeddings so far. How do the two dialogs fit together in a case of an embedding? It works because the second dialog fits into the first, and continues to move it forward towards its goal, despite the shift. But why are some shifts based on embeddings while others are not? It seems that there is no answer to this question so far. All that is known is that in some cases the one dialog fits in with the other while in other cases the two do not fit. What is known is that many of the informal fallacies are associated with illicit shifts where the transition from one dialog to another is not an embedding, and may be deceptive or unilateral.

The method of the investigation in chapter 6 is to begin with several cases of dialectical shifts. These cases are fabricated dialogs but the kinds of argumentation they represent are familiar enough. Some of the cases are extensions of interesting cases already known in the literature on informal fallacies. In other cases, there is an embedding of the second dialog into the first. In some of the cases, there is no embedding underlying the dialectical shift. In many of the cases, forms of argu-

ment that have a known argumentation scheme are used right at or around the shift. To assist the reader, these schemes are presented just after the presentation of the cases. Then each case is analyzed to study how the shift works. By comparing the various cases, and trying to see how the transition works in relation to the structures of dialog and argumentation in each case, the analysis moves ahead towards the proposed solution to the problem of embedding.

1   Dialectical shifts and fallacies

Dialectical shifts can be shown to be very important for dialog theory as a tool for analyzing argumentation, not only because they are very common in argumentation, but also because they are needed to explain how some of the most important kinds of informal fallacies work (Walton and Krabbe, 1995; van Laar, 2003). This claim can be easily appreciated by examining three fallacies, *argumentum ad baculum*, *argumentum ad consequentiam* and *argumentum ad ignorantiam*.

Many of the examples of the fallacious *argumentum ad baculum* given in the logic textbooks involve an overt or covert threat that is a species of argument from negative consequences. Whether such *ad baculum* arguments are fallacious depends on the context of dialog in which the argument has been put forward as a move in argumentation. It depends on the type of dialog the argument is supposed to be part of. In interest-based bargaining (negotiation dialog), threats are a fairly normal part of the argumentation strategies, whereas in a critical discussion, threats stand out as being highly inappropriate. In a critical discussion, the participants are supposed to present relevant evidence to support their viewpoints, and they are not supposed to pressure the other party to prevent him from presenting such evidence. The making of a threat contravenes this rule of a critical discussion constituting an instance of the *ad baculum* fallacy. What is revealed in such cases is that the very same argument, an appeal to threat, can be quite reasonable in one type of dialog, but can be fallacious if the dialog has shifted to a different type. Proving that such an argument is fallacious is a dialectical matter.

Walton and Krabbe (1995, 104) have made this identification of a fallacy based on a dialectical shift more explicit by introducing the fallacy of bargaining. The following example is given (p. 104): when a socialist minister of finance has been accused of taking advantage of questionable tax exemptions, he proposes that if his critics would abstain from moving for penalties for these exemptions, he will refrain from opposing a bill that these critics would profit from. The fallacy of bargaining has been committed here because there has been an illicit shift from the original persuasion dialog to a negotiation dialog. This was supposed to be a persuasion dialog to start with, but when the minister made the offer to refrain

from opposing the bill if his critics would abstain from moving for penalties for these exemptions, he shifted to a bargaining type of dialog. According to (Walton and Krabbe, 1995, p. 104) shifts from persuasion dialog to negotiation dialog are not always fallacious. However, in this case, the minister has evaded the issue by illicitly abandoning the original persuasion dialog. The shift is illicit because the negotiation dialog does not support the original persuasion dialog, and could even prevent it from continuing in a more productive manner towards its goal.

Argument from consequences is a reasonable kind of argument in which the proponent argues that a certain proposed action would be good or bad because good or bad consequences would follow from it. However, logic textbooks have sometimes cited argument from consequences as a species of fallacy. The classic cases are the following two examples from (Rescher, 1964 p. 82).

> Vegetarianism is an injurious and unhealthy practice. For if all people were vegetarians, the economy would be seriously affected, and many people would be thrown out of work.

> The United States had justice on its side in waging the Mexican war of 1848. To question this is unpatriotic, and would give comfort to our enemies by promoting the cause of defeatism.

It is not too hard to see why these examples are properly classified by Rescher as instances of the fallacy of arguing from consequences. In the first example, the issue is whether vegetarianism is an injurious and unhealthy practice. The kind of evidence that would be relevant to this issue would be an argument like one saying that vegetarians do not get enough protein, and that is unhealthy. However, although it may be true that if all people were vegetarians, that would have negative consequences on the economy, that is not relevant evidence to prove or disprove the claim that vegetarianism is unhealthy as a diet.

In the other case, the issue is a historical one on which country had justice on its side in the Mexican war. To answer this question, historians have to look at evidence of who started the war and so forth. Now it may be true at a particular time in history that to question this would be unpatriotic and would have negative consequences in promoting the cause of defeatism. That may be a practical reason for arguing that one should not question that the United States had justice on its side in these circumstances, but it is not a relevant reason for arguing that the United States had justice on its side, or did not. In both these textbook cases we can see that there has been a dialectical shift from a persuasion dialog about an issue to a practical deliberation concerning consequences of actions, and the negative impact of these actions. To understand that argument from consequences, a normally reasonable and very common kind of argument, is fallacious in these two cases, one has to grasp that there has been a dialectical shift from a critical

discussion type of dialog to a deliberation about how to act given certain conse-
quences that may be likely. This fallacy is a subtle one, because argument from
consequences is an inherently reasonable type of argument. To see that such an
argument is fallacious in these two cases one has to be aware that there has been a
dialectical shift, and it is the shift from the one type of dialog to the other that
makes the argument fallacious.

The third type of fallacy is that of the argument from ignorance. It can some-
times be a reasonable kind of argument, but in other instances is fallacious. An
example to illustrate how the *argumentum ad ignorantiam* can be reasonable is the
foreign spy case (Walton, 1989, p. 107).

> Mr. *X* has never been found guilty of breaches of security, or of any connection
> with agents of the foreign country he is supposedly spying for, even though the
> Security Service has checked his record. Therefore, Mr. *X* is not a foreign spy.

It is impossible to be absolutely certain beyond all doubt that Mr. *X* is not a foreign
spy for he could have had deep cover, and have destroyed all evidence that might
have been used to reveal his espionage. But suppose that a competent professional
security agency has done a search that turned up no evidence at all that Mr. *X* is a
foreign spy. Such negative evidence does support the conclusion that Mr. *X* is not a
foreign spy. The reason is the implicit premise that if Mr. *X* was a foreign spy, the
search by the agency would have discovered some evidence of his being a foreign
spy. Because this conditional premise cannot be proven beyond all doubt, however,
argument from ignorance is usually best seen as a defeasible argument at some stage
of a dialog or investigation in which evidence is being collected and assessed.

In everyday reasoning, inferences can be drawn not only from information
that has been collected as evidence, but also from failure to find information. Such
an inference is often called a lack of evidence inference. For example, suppose I am
in a shed that has a tin roof but no windows. I can't see outside, so I can't tell
whether it is raining or not. But suppose there is no noise, of a kind that would be
evident if it were raining outside. The sound of even the least bit of rain would be
amplified as it hits the tin roof, making it clearly audible to anyone inside the shed.
Based on the absence of such a noise, I can draw the conclusion that it is not rain-
ing outside. Of course, this is just a conjecture. If I could go outside and look, that
would be better evidence of what the weather is like. But still, the conjecture that
it is not raining is a conclusion I can draw by a lack of evidence inference. It can be
a reasonable conclusion to draw.

Lack of evidence inferences are very common in history, where they are called
ex silentio arguments. For example, if no evidence is found on tombstones or his-
torical records that the Romans awarded military medals posthumously, then it is
reasonable to conclude that they probably didn't. Lack of evidence inferences are

also very common in computing, when drawing inferences from a data base. In an expert system, knowledge in a domain of expertise is collected in a set of facts and rules called a knowledge base. When a user asks the expert system a question, it searches around in the knowledge base. If the item of information is there, the system presents it as the answer to the question. But if the information is not there, that finding too can be significant for answering a question. For example, an expert system on industrial and agricultural products in South America called SCHOLAR[1] is asked the question, "Is Guyana a major rubber producer in South America?". SCHOLAR searches around but does not find the statement 'Guyana is a major rubber producer' in its knowledge base. All it finds as the outcome of its complete search on major rubber producers in South America is that Peru and Columbia are major rubber producers. SCHOLAR answers the question with the following statement: "I know enough that I am inclined to say that Guyana is not a major rubber producer in South America". This example shows how a lack of evidence inference can be a reasonable argument.

Lack of evidence inferences have traditionally been treated as having a lesser standing in scientific research than positive results found in a study or experiment. For example, suppose a drug is tested for a specific toxic effect, like production of cancer cells. Let's say the drug is tested on mice and the mice are found to have cancer. This is a positive finding, and is called positive evidence. But suppose the finding of the experiment is that the mice exhibit no evidence of cancer after having taken the drug during the trial. This is a negative finding. But is it evidence of anything at all? Should the finding be published in a scientific journal? The questions are controversial. Some say that it would lead to a deterioration of standards of quality of scientific research if negative evidence were to be considered publishable. Others say that even a negative finding is a significant outcome that represents a genuine kind of evidence, even if it is negative evidence.

There is a longstanding tradition in logic that lack of evidence inferences are fallacious (Krabbe, 1995). This form of argument is traditionally called the argumentum ad ignorantiam, or argument from ignorance (literally, the argument to ignorance). A traditional example is the argument that extra-sensory perception must exist because all attempts to prove conclusively that it does not exist have failed. What the example may show is that when arguing about phenomena on the fringes of science where it is hard to define exactly what counts as evidence either for or against a claim, one must be careful about drawing any inferences, whether positive or negative. Even so, the argument from ignorance, or lack of knowledge

---

1.   Allan Collins, Eleanor H. Warnock, Nelleke Aiello and Mark L. Miller, 'Reasoning from Incomplete Knowledge,' *Representation and Understanding: Studies in Cognitive Science*, ed. Daniel G. Bobrow and Allan Collins, New York, Academic Press, 1975, 383-415.

inference, or ex silentio argument, is often a reasonable one, even if it is typically not highly conclusive. How strong or conclusive it is depends on how complete or incomplete the knowledge base is that served as the information on which the inference was drawn, and on how thorough the search through it was.

So conceived, argument from ignorance has two premises in the argumentation scheme below (Walton, 1996, p. 254).

Argumentation scheme for argument from ignorance

> Lack of Knowledge Premise: Proposition A is not known to be true (false).
>
> Conditional Premise: If A were true (false), then A would be known to be true (false).
>
> Conclusion: Therefore A is false (true).

The conditional premise depends on how complete the base is, in a given case, and how thorough the search through it has been. For example, in the case of SCHOLAR, we could assume that its knowledge about rubber producers in South America is, if not complete, at least quite deep. And we could assume that, since it is an automated expert system, its search through the knowledge base would be very thorough. If Guyana were a major producer, you can bet that SCHOLAR would know that. Since SCHOLAR answered the question by saying it does not know that, you can draw the conclusion that, based on SCHOLAR's expertise, Guyana is not a major rubber producer. Of course, this is only an inference. You yourself do not have the facts directly, because you are not an expert on rubber producers in South America. But under the circumstances, it can be a reasonable inference to draw.

The fallacious type of argument from ignorance is one where no serious attempt has been made to prove the conditional premise, and the arguer still tries to float the argument on the basis of little or no evidence. But in attempting to judge whether such an argument is reasonable or fallacious, we can often get into serious disputes about burden of proof, where each side tries to shift the burden to the other side. The dialog format of such a dispute as follows has been identified by Krabbe (1995, p. 256).

Burden of proof shift sequence

> Proponent: Why *A*? How can you prove it?
>
> Respondent: Why not-*A*? How can you disprove *A*?

This pattern of shifting back and forth of the burden of proof in dialog is characteristic of the more problematic cases of argument from ignorance.

Prakken (2001) showed how shifts in a burden of proof work in legal reasoning of the most common sort. He showed that such problems can only be solved by turning to meta-argument level considerations, using the following example (p. 259) summarized below.

> Plaintiff supports his claim that a contract exists by arguing that there was an offer and acceptance by the defendant. She supports this claim by bringing forward two witnesses who testify to her offer and defendant's acceptance. The burden is now on the defendant to question or refute this evidence. Defendant attacks her argument by presenting evidence that the witnesses are unreliable. The burden now shifts to his side, or risk refutation of the argument based on witness testimony. How should the issue of which side bears the burden of proof be decided? Should defendant have to prove his claim that the witnesses are unreliable? Or should plaintiff have to prove the opposite proposition that the witnesses are reliable. On which side should the burden to prove or refute this claim lie?

This common example shows how a burden of proof can shift back and forth, and why ruling on which side should have the burden of proof may require a decision by a third party, a judge. A metadialog is a dialog about the dialog (Krabbe, 2003, p. 83). In some cases, like that of a dispute about which side has the burden of proof, the original dialog may move to a metadialog in order to have a secondary dialog on how the moves in the first dialog should be judged (Krabbe 2003). Hard cases can require systematic intervention, and cannot be fairly adjudicated without the intervention of a third party at a metadialog level. In such cases, there needs to be a shift from the original dialog to a metadialog, and then back to the original dialog. According to Wooldridge, McBurney and Parsons (2005), argumentation in dialog is inherently metalogical, meaning that it does not just involve the asserting of statements about some domain of discourse, and putting forward arguments based on these statements, but also the making of arguments about these arguments at a higher level.

Krabbe (2003, p. 83) showed that many dialog moves may ask for conversational repairs of some sort that could prevent blockage of the dialog from achieving further progress at that level. A burden of proof impasse where one side says "You prove it!" and the other side replies "You disprove it" is a case in point. Trying to resolve such a dispute within a dialog, like a persuasion dialog for example, might prove futile. Hence one very important type of dialectical shift is the shift from a dialog to a metadialog.

2 The problem of shifts and embeddings

The problem studied next is to find a structure that reveals how one dialog fits into another in a sequence of argumentation where one type of dialog is superseded by another. The problem is to find a method for analyzing and evaluating cases of argumentation where a dialectical shift has occurred. There are several different kinds of dialectical shifts studied in (Walton and Krabbe, 1995, pp. 100–116). One common kind of example that is studied here is the shift from deliberation dialog to negotiation dialog. Suppose for example that a group of executives are having a discussion in a business meeting in a conference room about whether to launch a new product. They are collectively deliberating on which course of action to take. But once the decision is made to go ahead with the new product, and everyone has agreed that launching the new product is the action that the company should take, the deliberation on that issue is at an end. The discussion may then shift to other matters. Having decided to go ahead and market the new product, they may start to negotiate on who should take responsibility for each part of the task of pricing and marketing the product. The argumentation in the two dialogs may be interconnected. But there has been a definite shift from the one type of dialog to the other. And to properly analyze and evaluate the sequence of argumentation, taking the dialectical shift into account may be vitally important. Attempting to evaluate the argument without paying attention to the context of its use, and how that context has shifted, could be an over-simplification that results in a superficial analysis.

The concern with dialectical shifts arises out of recent developments in argumentation theory, especially informal fallacies, as noted above, and builds on some of the methods that have been developed to this point. The construction of systematic structures for different types of dialog that represent frameworks of argument use has been carried out by several investigators in recent years. Several publications that should be cited are Hamblin (1970; 1971), Rescher (1977), Hintikka (1979; 1992; 1993; 1995), Mackenzie (1981; 1990), Barth and Krabbe (1982), Hintikka and Hintikka (1982), Carlson (1983), and Walton and Krabbe (1995). These works refer to others too numerous to be listed here. The development of this dialectical approach to logic derives from two independent initiatives. One is the Lorenzen School, whose exponents modeled logical reasoning as an orderly sequence of moves in a question-reply sequence (Lorenzen, 1969). The other is the construction of different dialectical systems, or formal models of dialog by Hamblin (1970; 1971) to study logical fallacies. Another formal structure that will be shown to be applicable is the use of directed graphs as argument diagrams to model argumentation chains. The method of argument diagramming is widely in use, but its structure has been analyzed in (Freeman, 1991). A structure that models a chain of argumentation as a digraph has been presented in (Walton, 1996a). Using

this structure, a chain of argumentation in a given case can be modeled using a directed graph where the vertices are the premises and conclusions of what is taken to be the given argument.

The third historical root of the dialectical approach to the analysis of argumentation is the famous paper of Grice (1975) on the logic of conversation. According to Grice, an argument, or move in argumentation, should be seen as a collaborative contribution to a conversation between two parties. According to Grice's pragmatic approach, an argument should be evaluated on the basis of how well it contributed to the moving forward of a conversation towards its goal, at the stage of the conversation where it was put forward. This pragmatic approach advocated judging an argument with respect to how it was used for different purposes in different types of conversational exchanges. Grice's pragmatic analysis offered the beginnings of a framework of rationality for judging an argument with respect to how it had been used in a given case to contribute to goals of dialog appropriate for the case. But what are these types of dialog? To refer to Table 1.2 in chapter 1 again, these six basic types are persuasion dialog, the inquiry, negotiation dialog, information-seeking dialog, deliberation, and eristic (quarrelsome) dialog.

To put the problem in perspective, let us elaborate a bit on the properties of persuasion dialog, drawing on the exposition of the critical discussion type of dialog in chapter 1. In persuasion dialog, there are two types of opposition representing two kinds of cases. In a case of weak opposition, the proponent has a particular thesis to be proved, and the respondent has the job of casting doubt on that thesis by raising critical questions. In a case of strong opposition, each participant has a thesis to be proved, the one thesis is the opposite (negation) of the other, and each participant has the aim of persuading the other to accept his or her thesis. In a persuasion dialog, each party takes the commitments of the other as premises, and then by a series of steps in a chain of argumentation, tries to persuade the other party to come around to giving up her or his doubts. According to the system of classification advocated in (Walton, 1998) and followed in chapter 1, the critical discussion type of dialog is a subspecies of persuasion dialog. The goal of the critical discussion is to resolve a conflict of opinions, meaning that one side is proved right and the other side is proved wrong. But according to the account given in (Walton, 1998), persuasion dialog can be successful in some cases if there were improvements in formulations of the arguments on both sides, even if the conflict was not resolved. Once persuasion dialog has been recognized as a distinctive type of dialog in its own right, it is easy to appreciate that argumentation in the various other types of dialog is different in some ways, and hence that there can be shifts from one type of dialog to another. Further description of the goals, rules and moves for each type of dialog can be found in (Walton and Krabbe, 1995). Formal dialectical systems representing the inquiry and information-seeking types of dia-

log have been presented by Hintikka (1979; 1992; 1993; 1995). All of this is by way of review, but it is necessary to be clear about it, in order to grasp the significance of the concept of the dialectical shift.

At the level of formal dialectical systems, the problem of dialectical shifts at first appears to be simple. Each formal dialog is a sequence of moves in which the two participants take turns making moves of specified kinds (Hamblin, 1971). The dialog rules govern what kinds of moves may be made. For example, the proponent may be allowed to start off a dialog by asking a question. Then a dialog rule will specify how the respondent should make the next move. For example, he may be allowed to either give an answer, or to say "I don't know." Or in addition to one of these replies, he may be allowed to ask another question. So if there is a shift, and one party starts up a different type of dialog, then any such move would simply violate the dialog rules for the initial type of dialog they were supposed to be engaging in. Such a breaking of the rules presumably would indicate a dialectical shift. However the problem is not so simple. In some cases, the kinds of move permissible in one type of dialog are the same as the kinds of move permissible in another type of dialog. So there may be a shift even though the move appears (in light of the dialog rules) to be appropriate. This kind of situation will be illustrated in the cases studied below. Also, sometimes a shift is a good thing, because it is based on an embedding. A purely syntactical comparison of moves over a shift would not distinguish between embeddings and cases of shifts where there is no embedding. And this is the very distinction that is vital to make.

Reed (1998) has presented a way of formalizing sequences of argumentation in a dialog in which dialectical shifts can be marked in the notation. Reed's method of formalization follows the pattern of Hamblin's, in which a dialog is seen as an ordered sequence of moves. A dialog frame (Reed, 1998, p. 248) is a four-tuple composed of a type of dialog, a topic, a pair of participants, and a sequence of utterances. The type of dialog, t, can be persuasion, negotiation, inquiry, deliberation or information-seeking. The topic, τ, is the original issue or basic conflict that is supposed to be resolved by the dialog. The two participants in a dialog are x_i and y_i. Each utterance made at each move of the dialog is numbered. So $ux_i \rightarrow y_i$ refers to the ith utterance in a dialog between the two participants. In the simple system presented by Reed for modeling dialectical shifts, each utterance has the form of a pair in which the first element is a statement and the second element is a support for that statement. A dialog frame F is defined as follows (p. 248).

$$F = \langle \langle t, \Delta \rangle \in D, \tau \in \Delta, (u^0 \, x_0 \rightarrow y_0 \,, ..., u^n \, x_n \rightarrow y_n) \rangle$$

In Reed's notation, a dialog is a sequence of moves in which the participants take turns making utterances of the permitted types. For example, in a negotiation dia-

log, proposals can be made and then the other party can accept that proposal or not. The following very simple case is used by Reed to show how his system for modeling dialectical shifts works (Parsons and Jennings, 1997, p. 267). Two agents are engaged in deliberation dialog on how to hang a picture. They realize they need a hammer. One knows where a hammer can be found, and then the two begin to negotiate on who will go and get the hammer. In Reed's notation, the type of dialog is represented at each move. So, in this case, one move would be marked as "deliberation" and then just after the shift, the next move would be marked as "negotiation". The formalization of the dialog would not only include the type of move and refer to the participant who made the move, as in Hamblin's system. It would also include reference to the type of dialog the move was part of.

Although Reed's system presents us with a method of modeling dialectical shifts in formal dialectic, it leaves open the problem of determining when such a shift is an embedding as opposed to a shift that is not based on an embedding. It also leaves open the problem of determining in any actual case of natural language argumentation, when such a shift can be detected. Reed's system of formal dialectic makes the problem of embedding much more interesting, because it offers a framework in which shifts can at least be modeled formally within the representation of a dialog. But how do embeddings actually work? What is the structure over the dialectical shift that fits the two adjacent dialogs together into an embedding?

The above outline has sketched out the existing methods currently used in argumentation theory that are relevant to the problem of finding out how embeddings work. These existing methods need to be extended to solve the problem of embeddings, and thereby to cope with various problems known to be posed by cases of dialectical shifts. In some cases the dialectical shift is associated with problems and fallacies (Walton and Krabbe, 1995; Walton, 1998). The shift represents a deterioration or even blocking of the proper progress of the argumentation. In other cases, the shift seems to represent an improvement. The advent of the second dialog actually appears to be a constructive development. Typical of the latter cases is a fit, or what is called a functional embedding of the one dialog into the other. In some cases, there can even be a so-called cascading or sequential series of shifts among several types of dialog (Walton and Krabbe, 1995, p. 107). The problem set for this investigation is to grasp how functional embeddings work, and to reveal the structure of the embedding whereby the two (or more) dialogs fit together. How is it that a sequence of argumentation can continue over a dialectical shift in a constructive way based on a functional embedding? How can the embedding surrounding the argument be modeled as having a systematic structure? How can such an embedding be identified, analyzed and evaluated in a given case? Can examination of some common kinds of cases of shifts reveal significant insights on how to recognize dialectical shifts and pinpoint the location of the transition from

the one dialog to the other? What local features need to be recognized as tying with and indicating shifts. What clues do these features yield, and how can they be used as evidence from the text of discourse in a case to show that a shift has occurred in a case and to justify the determination of a shift?

3   Cases of shifts based on embeddings

In this section, several cases of dialectical shifts are presented. These cases are not actual cases from transcripts of real discourse. But they are realistic enough so that the reader can easily relate them to his or her own common experiences of everyday practices of argumentation. Some of the cases are new, while others are extensions of cases that are already known in the literature. In each of the cases in sections 1 and 3, there is a dialectical shift from one type of dialog to another. But only the shifts in the cases in section 3 are embeddings.

The Fence Case
Bob and Ed are neighbors. Bob has decided to put up a new fence. But one side of his yard borders on Ed's property. Bob and Ed agree that they will share the cost and labor of this common section of fence. Then they deliberate together on how they will build this part of the fence. They discuss what kind of wood they will use, what it will cost, how they will paint the fence, and so forth. They agree on everything, except on the question of how the post holes should be dug. Bob says that he will put in the holes for the posts, and he and Ed can dig together on the holes in their shared part of the fence. But Ed says that he has a bad back, and can't help dig post holes. But he has a cousin Norm who will dig the holes for ten dollars a hole. That seems expensive to Bob. Bob plans to dig his own post holes on his part of the fence, but they disagree on how the post holes should be done on the shared part of the fence. This part of the dialog runs as follows.

> Bob: We'll need five post holes on the shared part of the fence. We can easily dig these holes together. It won't take very long.
>
> Ed: I have a bad back. My cousin Norm will dig these holes for ten dollars a hole. This is an easier way to do it.
>
> Bob: Ten dollars a hole is too much. That means I have to pay five dollars for each hole. But I can easily dig them with the post hole driller I will already have to rent.
>
> Ed: Well, suppose I pay five dollars to you for each hole in the common part of the fence, and then you can dig the holes. That way I don't have to risk back problems.

> Bob: OK. It's a deal.

In this case, Bob and Ed started out in a deliberation dialog on the best way to build a new fence along their common property line. They discussed how high the fence should be, where it should be along the property line, and so forth. It began to seem like the project was feasible. Then they ran into the problem of the post holes. There was disagreement here on how to proceed. Then, in the sequence quoted above, the exchange shifted to a negotiation dialog. They began to negotiate on how they would share the costs or labor for the digging of the post holes.

The argumentation in this case shifted from a deliberation dialog to a negotiation dialog. But was the sequence of negotiation argumentation quoted above relevant? Yes, it was, because it fitted constructively into the prior deliberation dialog. Once the deal about the post holes was successfully concluded, the problem posed in the prior attempt at deliberation was solved. Bob and Ed could then agree on a joint plan for building the fence. Bob could then go ahead and order the lumber, and plans for building the fence could proceed to the next stage. So the negotiation dialog contributed to the deliberation dialog. Indeed, it was an essential part of it. The deliberation dialog would not have successfully resulted in a joint plan of action if agreement on the post hole issue had not been secured by negotiation.

The PC Case

Bill and Edith needed a new personal computer for a research project they were working on together. The old PC they were using had crashed, and it was no longer on warranty. Getting it serviced might be very expensive, and their budget for hardware was limited. But they needed a PC right away. They had to take action. But they were confronted with a lot of confusing questions. What features do they need? Do they need to have a modem on the machine? Does it have to be a particular type of modem? How much memory should the PC have? What type of monitor should they get – a cheap one, or a more expensive one that might not be so hard on the eyes? Should they get a laptop computer or a desktop model? Here Bill and Edith were confronted with a practical problem that required action, but neither of them knew very much about technical requirements and features of recent personal computers. Although they discussed the problem, they had a hard time knowing which PC they should get. They turned to their younger colleague Brent for advice. Brent knew a lot about recent PC's, having just bought one himself. Brent was experienced with computers.

> Bill: Even if we can get the computer fixed, it's not reliable. I think we need to get a new one.

> Edith: Yes, I think so too. But what kind should we get? The budget is limited.

Bill: The laptops are more expensive. How has your new desktop computer worked out for you, Brent?

Brent: It has worked out fine for me. I have had no serious problems with it, and the help I got with minor problems under the warranty agreement was very good.

Edith: Here is the list of all the features on the machine we were thinking of buying. There is no modem listed. Also, there is no ethernet card listed. Do you think we will need these features, since we plan to use the PC exclusively at the office?

Brent: You will need the ethernet card, because they do not provide one here. You will only need the modem if you plan to use the PC away from the office.

Bill: Are there any other features we will need that are not on the list?

Brent: No, this machine has all the features you will need.

In this case, as Bill and Edith ask Brent questions about computers, they are engaging in an information-seeking dialog with Brent. The problem confronted by Bill and Edith was one of deliberation. They needed a PC, and were in the position of having to take action soon. But they had heard that some people who had bought new PC's at the office had experienced a lot of problems with them. They had already had problems with the old PC they were using on the project, and wanted to avoid such problems in the future if they could. Hence the information provided by Brent proved to be very valuable for helping Bill and Edith to arrive at a decision they felt would be right for their project.

In this case, there was a shift from a deliberation dialog to an information-seeking dialog. But the information-seeking exchange was relevant to the original deliberation dialog. Why? It was relevant because it helped Bill and Edith to arrive at an intelligent choice on the problem of which PC to buy. They were able to make a better informed decision because of the dialog they had with Brent. So the information-seeking dialog contributed to the deliberation dialog. The second dialog fitted constructively into the first.

The Common Currency Case
Helen and Roger were having a critical discussion in 1980 on the issue of whether adopting a common European currency would improve the financial situation in Europe. Neither was an expert in economics, and they agreed that collecting some relevant information about the economic facts about common currency situations would improve the critical discussion. Many of the arguments on both sides seemed to turn on factual questions – for example, on how well such currency

conversions have worked in the past. Many of the arguments on both sides of the critical discussion took the form of argumentation from consequences. Helen argued that conversion to a single currency would have bad consequences, and therefore it is a bad policy. Roger argued that conversion to a single currency will have good consequences, and therefore it is a good policy. But the question of what consequences might be expected to occur is based on many factual considerations of the kind that are known to economists. Helen and Roger agreed that the critical discussion would be more interesting if it were based on what is known about outcomes of adopting a common currency. Part of the dialog ran as follows.

> Helen: Conversion to a common currency will have the bad outcome that if even one single country has economic problems, it could be a disaster for all Europe. And of course, as we all know, one country will have economic problems.
>
> Roger: Well yes, that could be a problem to be concerned about. But if one country is having problems, the impact of these problems could be lessened by a common currency. Also, a common currency would mean that Europe will be able to compete in world markets on an equal footing with North America. This will lead to a much more stable European economy.
>
> Helen: Well, stability can be a good thing, but it can also lead to bad things.
>
> Roger: Dr. Smith, a professor of economics at Harvard, has studied many cases of common currency conversion among neighboring countries.
>
> Helen: Well yes, I agree that Dr. Smith is an expert on currency conversion.
>
> Roger: Dr. Smith says that conversion to a common currency improved the economies of all participating countries in cases where it has been adopted.
>
> Helen: What evidence did Dr. Smith base that conclusion on? How many cases of past currency conversion did she study? Which countries were involved? How long ago did these conversions take place?
>
> Roger: One case she studied was that of Bosnia in the nineteenth century.
>
> Helen: Well, I doubt that any of these past cases is similar to the case of a common currency for all of Europe.

To provide justification for his line of argument, Roger quoted an expert source, the economist Dr. Smith. Helen agreed that Dr. Smith was an expert in the field of economics who had studied currency conversion. Using an appeal to expert opinion, Roger then put forward an argument. This argument will be analyzed below, where it will be called argument alpha. What will be shown below is that the dia-

lectical shift occurs right within argument alpha. A key premise in argument alpha comes in to the critical discussion from what is presumed to be a prior expert opinion consultation dialog – a species of information-seeking dialog.

The Euthanasia Case

Wilma and Bruce are having a discussion in their ethics seminar on the issue of physician-assisted suicide. Bruce advocates the thesis that permitting physician-assisted suicide is a good policy idea. Wilma is against it.

> Wilma: Once you allow this, the problem is that it is impossible to confine it to voluntary cases. Physicians will kill people against their will to get spare body parts. Disrespect for life will set in. Soon, any persons deemed socially unfit or undesirable will be euthanized. The outcome in the end will be mass exterminations, like in the holocaust.

> Bruce: Wait a minute. How do you know one thing will lead to the other?

> Wilma: Because you can't confine any program like this to voluntary cases. Many terminally ill patients are not autonomous decision-makers. The next of kin and the physician can have an undue influence that can be exploited. Then such practices of exterminating bothersome patients become accepted.

> Bruce: But surely you can have laws that state the individual's right to choose and that see to it that this right is not violated.

> Wilma: No way. Once you have any policy like this, that permits physicians to kill, it inevitably gets violated and expanded.

> Bruce: Well, I'm not so sure about that. I have heard that in some countries, they already have programs of physician-assisted euthanasia for terminally ill patients who want to avoid needless suffering. From what I have heard, these programs have been successful in confining euthanasia to those who have voluntarily requested it from a physician.

> Wilma: Where did you hear that?

> Bruce: I have an article on it right here. Here, I have some parts of it underlined. Let me read some of these parts.

In this case, Bruce goes on to read some parts of the article that describe what is currently happening with the program for voluntary physician-assisted euthanasia in Holland. Wilma asks how long this program has been going, and how it has changed since it started. Bruce reads out some parts of the article that describe the history of how the program was originally set up several years ago, and how it has progressed since then. Wilma asks whether the article says anything about how people in Holland have reacted to the program, and whether they are worried

about it being expanded beyond voluntary cases. Bruce reads parts of the article that report on these matters.

In this case, there was a dialectical shift from a critical discussion (persuasion) type of dialog on the euthanasia issue to an information-seeking dialog. The source of information was the article that Bruce read from. Originally the discussion was a persuasion dialog in which Wilma and Bruce disagreed on the issue of physician assisted suicide. Then it shifted to an information-seeking dialog in which Bruce reported some information he had found on developments in Holland. The argumentation in the dialog continued right over the shift.

An interesting question is whether the argumentation after the shift to the information-seeking dialog should be judged to be relevant to the argumentation in the critical discussion. It is relevant, as can be seen by examining the case. Wilma put forward a slippery slope argument. The gist of it that once you start allowing physician-assisted suicide, you can't stop expansion of the killing until eventually you reach a holocaust situation. How could Bruce criticize or refute this argument? There could be many ways. But suppose he could find an actual case where a policy of physician-assisted euthanasia had been in place for a while, and the supposed abuses did not occur. This information would be relevant, because it counts as evidence against the plausibility of Wilma's slippery slope argument. Hence the information-seeking dialog, in which Bruce read excerpts from the article to Wilma, was relevant.

4   Cases of shifts not based on embeddings

The Bicycle Case
In this case, Karen and Doug were cycling along the bicycle path having a critical discussion on the pros and cons of living in a house versus a condominium. During the conversation they came to a fork in the bike path, with signs indicating which town each path goes to. They stopped and had a discussion about which way to go. In this case the shift was necessitated by a practical need to make a decision. As they cycled along, Karen and Doug were initially having a critical discussion about the condominium issue. But then there arose the need to engage in a deliberation type of dialog on which path to take. Once a decision was made, and the deliberation was concluded, they could then resume the critical discussion. In this case, the deliberations stopped the critical discussion temporarily. But this discussion could easily be taken up again, once the decision was made on which path to take.

Doug: One disadvantage of a condominium is that you can sometimes hear the neighbors.

Karen: Yes, but an advantage is that you have more security if you travel a lot.

Doug: But then there's that "common element" fee. That can go up.

Karen: OK. There's the fork in the road. Which way do you want to go, to Katweg or Lisse?

Doug: Well. Lisse is quite a ways. On the other hand, we went to Katweg last week.

Karen: We haven't been to Lisse, and I hear it's very nice.

Doug: OK. Let's take the path to Lisse (They start riding down the path again).

Karen: Well, yes, the common element fees can go up, but house taxes can also go up.

Doug: Yes, you are right. House taxes get higher every year.

In this case the dialectical shift could be described as an interruption. First Karen and Doug were having a critical discussion on the condominium versus house issue. Then the need to make a prompt decision intervened.

The Airline Case

A major airline, suffering a huge deficit, was forced to cut costs. They hired a new chief executive who was known for fiscal prudence. The union did not like the new executive, Frank Black, and saw him as a symbol of capitalist greed and ruthlessness. However, this portrayal of Frank Black was not really accurate. Many saw him as a quiet man who led a simple life, and was not especially materialistic. During negotiation talks between the union head, Bob White and the new executive, Frank Black, the following dialog occurred.

Black: If the union could just take a reasonable cut in salaries and benefits, we could make a go of it.

White: Management should take salary and benefits cuts as well.

Black: Well, yes. That's reasonable.

White: But we are not agreeing to reduction of overtime hours.

Black: You have to agree to that, because it's just costing too much to support, and its not good for anyone.

White: The workers need that overtime. We have families and the cost of living is high.

Black: Yes, I understand. But you just can't have it. The company can't afford it.

White: You see, that's your problem. You are against the ordinary working people.

Black: I am not.

White: You are a pillager of the American dream.

Black: You and your whole pack of morons are ruining the company by your obstructive and greedy tactics.

White: You are a slimeball.

In this case the dialog started out as a negotiation attempt, but then degenerated into a quarrel. The shift can first be detected precisely at the point where White accused black of being against the ordinary working people. This allegation accused Black of being one-sided and divisive. It accused him of being against the other side. The suggestion is that Black has a bad attitude, and that there is no point in trying to argue with him, because he will not be reasonable. Black simply denied the charge. Then White followed up by using the negative epithet, "pillager of the American dream", an expression that has all kinds of bad connotations, suggesting that Black is a bad person. At this point, Black rose to the bait, and replied with some equally vituperative epithets. At this point the dialog has turned into a quarrel.

In this case, the dialog gradually shifted from a negotiation to a quarrel. But there was one single point, as noted above where the quarrel was precipitated. Although the shift began to occur at a single identifiable point, it took place over a sequence of moves. Starting from this initial point, there was an escalation. First Black was said to be "against the ordinary working people." The next step was to call him a "pillager of the American dream". This was an even worse allegation. At the end, he was called a "slimeball". This was the most serious stage of the escalation. In this case then, the shift occurred at one definite point, but then the dialog gradually deteriorated from that point onwards.

The shift was not entirely unilateral. It was precipitated by the one party, but then eventually the other party joined in to the quarrel. It was a gradual shift that picked up momentum as it went along. Black held back at first. But then, after White called him a "pillager of the American dream", he replied by calling White and his fellow union member a "pack of morons". The airline case is a fictional extension of the real case cited in (Walton, 1992, p. 218) that happened in 1989. In the real case, the negotiation deadlock between the union and the chief executive officer led to the eventual bankruptcy of Eastern Airlines.

The infomercial case

This case is an extension of a real case cited in (Walton, 1998, p. 209). In a 30-minute TV Program called *Consumer Challenge*, two speakers, Dan and Zelda, are

shown, who look like news anchors sitting at a desk. The general appearance of the format looks to the viewer somewhat like that of *60 Minutes*. The two speakers look like reporters who are reporting on events shown in videotape clips interspersed in the program. In the introduction, Dan said that in the program, they would be presenting a consumer's evaluation of popular brands of sunglasses currently available at retail outlets.

Zelda: We conducted an objective evaluation of several available brands, and our staffers tested them to see which ones performed best.

Dan: Brand X caused me to squint when I looked into bright sunlight.

Zelda: Our investigator reported that Brand Y slipped down her nose constantly.

Dan: I tested out the Pilot sunglasses, and they were terrific. They not only stayed on my nose, but they did not cause squinting.

Zelda: These Pilot sunglasses really look great, for both men and women.

(Visual shows a male model and female model, both wearing Pilot sunglasses.)

Dan: They are priced very reasonably, considering their high quality.

Zelda: You can get them for $12.98 at all Topline Store outlets today.

Dan: Or you can just dial our toll-free number (shown on screen) and we can send you a pair right away, with no delivery charge.

Zelda: Get a pair right away. They are so cool.

In this case, the program started out by giving the appearance of being a kind of consumer report type of presentation giving an independent evaluation of several different kinds of sunglasses. The viewer was told that several brands had been tested, and that an objective report would be given on how the sunglasses performed in the tests. The format appears to the viewer to be one of presenting information. But then gradually a shift to a sales pitch took place. The sales pitch is a kind of advocacy dialog that is different in important respects from an objective investigation or evaluation of the kind we are familiar with from *Consumer Reports*. But by the end of the sequence of argumentation in the dialog above, it is clear that the dialog is now a sales pitch.

A hint is given of the direction the dialog is taking when Dan says Pilot sunglasses are "priced very reasonably, considering their high quality". But the move where Zelda says, "You can get them for $12.98 at all Topline Store outlets today." is the precise point in the dialog where the shift to the sales pitch occurred.

The Mexican war case

Many of the traditional informal fallacies cited in logic textbooks are failures of relevance. One fallacy of this type is the *argumentum ad consequentiam*, or argument from consequences. The following dialog is an extension of the example of the fallacy of *argumentum ad consequentiam* cited in (Rescher, 1964, p. 82). In the dialog, Harry and Louise are having a critical discussion of the issue of which side "had justice" on it in the Mexican war of 1848. Harry thinks that the United States had justice on its side, but Louise is not convinced by his arguments.

> Louise: How can you prove that the United States had justice on its side in waging the Mexican war of 1848? Wasn't it simply a dispute over territory in which both sides should take some of the blame?
>
> Harry: Mexico provoked the United States.
>
> Louise: How can you prove that claim? I don't know of any evidence for it all.
>
> Harry: How can you doubt that the United States had justice on its side? To question this is unpatriotic, and would give comfort to our enemies by promoting the cause of defeatism.

What is evident in this small segment of dialog is that there has been a shift that occurs as soon as Harry puts forward his argument about "giving comfort to our enemies". Before this point, the dialog was a critical discussion of a specific issue. But as soon as Harry brought forward his argument from consequences, citing the supposed negative consequences of Louise's doubting of his position, there is a shift to a practical type of discourse (deliberation). Now the issue has become the practical question of whether expressing doubt could be somehow wrong, and lead to bad consequences. But after the shift Harry's use of argument from consequences is not dialectically relevant. It really has no appropriate place in the critical discussion. The dialog was originally supposed to be a critical discussion in which both sides try to present rational arguments for or against Harry's claim. If the original dialog had been a deliberation on how to avoid defeat in a current war by avoiding defeatism, the argument from consequences would have been relevant. The argument seems to be weak and implausible, even when it is evaluated from the point of view of a practical deliberation about the dangers of appearing to promote defeatism. But when evaluated from the viewpoint of the critical discussion, it is not even relevant. What this case shows is that argumentation from consequences can be a perfectly reasonable argument in the right context of dialog, but if that context shifts it can become a fallacy.

5   Argumentation schemes

Before attempting to analyze these cases, it is necessary to review some facts about argumentation schemes that have already been presented in chapter 3, and to add some new facts about them. Much has been written about deductive and inductive forms of argument in logic, but many common arguments fall into a third category that could be called presumptive, abductive, conductive, etc. The terminology is not yet agreed upon. Some common forms of argument falling into this third category have been classified as argumentation schemes in (Hastings, 1963), (Kienpointner, 1992) and (Walton, 1996). Each argumentation scheme has a matching set of critical questions. If an argument is put forward in a dialog by a proponent, and it meets the requirements of the argumentation scheme, and the premises are acceptable, then a weight of presumptive acceptability is thrown onto the conclusion. If the respondent asks an appropriate critical question, however, that weight of acceptability is withdrawn, until the question is given a satisfactory answer by the proponent. Several argumentation schemes are important for evaluating the argumentation in the cases above.

The argument from position to know is a familiar form of argument in which one party asks a second party for information that the second party is presumed to possess. For example, suppose a stranger to a city asks a passerby where the Central Station is located. The stranger presumes that the passerby likely has this information, since the passerby looks like a person who is probably familiar with the area. If the passerby does not have the information, little is lost by asking. This argument has the following general form.

Argument from position to know

> Major Premise: Source a is in a position to know about things in a certain subject domain S containing proposition A.
>
> Minor Premise: a asserts that A (in Domain S) is true (false).
>
> Conclusion: A is true (false).

Matching the argument from position to know are the following three critical questions.

> CQ1: Is a in a position to know whether A is true (false)?
>
> CQ2: Is a an honest (trustworthy, reliable) source?
>
> CQ3: Did a assert that A is true (false)?

The argument from position to know is taken as a fallible kind of argument that shifts a burden of proof in a dialog towards one side on a balance. If one of the ap-

propriate critical questions is posed by the respondent, the burden shifts to the other side. If the question is answered satisfactorily by the proponent, then once again the argument from position to know goes forward.

Argument from expert opinion, sometimes also called the appeal to expert opinion, is a special subtype of argument from position to know. Argument from expert opinion has the following general form (argumentation scheme), as indicated in chapter 3, section 2.

Argument from expert opinion

> Major Premise: Source *E* is an expert in subject domain *S* containing proposition *A*.
>
> Minor Premise: *E* asserts that proposition *A* (in domain *S*) is true (false).
>
> Conclusion: *A* is true (false).

Appeal to expert opinion is a defeasible form of argument, because experts can be, and frequently are, wrong. Argument from appeal to expert opinion brings forward a weight of presumption in favor of the acceptability of the conclusion, on a balance of considerations. The asking of appropriate critical questions by the respondent in a dialog throws a weight against the argument, suspending whatever plausibility the argument had until the critical questions are answered.

According to (Walton, 1996, p. 76) the argumentation scheme for argument from consequences is the following scheme. The letter *A* stands for a proposition that can be made true (brought about) by an agent.

Argument from consequences

> Premise: if *A* is brought about, then good (bad) consequence will occur.
>
> Conclusion: *A* should (not) be brought about.

This scheme is a composite that can take two forms. In positive argument from consequences, it is argued that *A* should be brought about because bringing about *A* will have good consequences. In negative argument from consequences, it is argued that *A* should not be brought about because bringing about *A* will have bad consequences. The slippery slope is a composite type of argument that is built on argument from negative consequences. As shown in (Walton, 1996, pp. 95–110), there are four basic types of slippery slope argument.

The slippery slope argument in the euthanasia case is of the type called the full slippery slope argument. The full slippery slope is a composite of three more simple types of slippery slope called the *sorites*, the precedent, and the causal type (Govier, 1982). The full slippery slope type of argument is based on a sequence of cases C_0, C_1, …, C_n, in which each case is related to the next case in the sequence.

The relationship of each case to its next (neighboring) case is called an "individual linkage". The sequence represents a gradual transition from one step to the next. The argumentation scheme for the full slippery slope argument is shown as follows (Walton, 1992, pp. 199–200).

Full slippery slope argument

> Initial Premise: Case C_0 is tentatively acceptable as an initial presumption.
>
> Sequential Premise: There exists a series of cases C_0, C_1, …, C_{n-1}, where each step leads to the next by a combination of causal, precedent and/or analogy steps.
>
> Group Opinion Premise: There is a climate of social opinion such that once people come to accept each step as plausible, then they will also be led to accept the next step.
>
> Unacceptable Outcome Premise: C_{n-1} leads to an ultimate outcome C_n (the horrible outcome), which is not acceptable.
>
> Conclusion: C_0 is not acceptable (contrary to the presumption of the initial premise).

Matching the full slippery slope are the following critical questions (Walton, 1992, pp. 201–206).

> CQ1: How strong is the evidential support for each of the individual linkages?
>
> CQ2: How strong does the argument need to be, in order to fulfill its burden of proof?
>
> CQ3: Can the projected progression be stopped by a firm boundary or rule?

Full slippery slope arguments are quite complex. Each individual linkage in the chain can be challenged. Many of these links are based on empirical or causal claims, or predictions about what will happen in the future. All such claims require evidence. Thus there is much scope for asking the first type of critical question. How much evidence is required to support a full slippery slope argument depends on how strongly the conclusion is worded. Does it say that the horrible outcome must occur, or that it will occur, or only that it might occur? Especially hard to support are cases where strong language using terms like "inevitably" or "unavoidably" is included in the conclusion.

The *argumentum ad hominem* is use of personal attack by one party (the proponent) in a dialog to try to refute or attack an argument put forward by the other party. Thus not all personal attack arguments should rightly be classified as *ad hominem* arguments. There must be the right dialog structure involved, where the party who is attacked (the respondent) has put forward a prior argument. But in

its most typical form, the *ad hominem* argument also has to be situated in a dialog in which three participants are involved – a proponent, a respondent, and an audience. However, in some cases, the audience and the respondent could be the same person. In the argumentation scheme presented below, the audience, or third party, is the judge of credibility. This form of argument is based on what is called a credibility function. A credibility function takes as its input value the credibility of the respondent and takes as its output value the plausibility of that respondent's argument. As the input value is lowered or raised, the output value is lowered or raised accordingly. So if an arguer is persuasively attacked as a bad person, and the plausibility of his argument depends on his credibility in a dialog, the plausibility of his argument will be reduced. The *argumentum ad hominem* has various forms (Walton 1998), but the basic form is the direct variant, represented by the argumentation scheme below.

Direct *Ad hominem* argument

> Major Premise: If the respondent is not credible, then his argument should not be judged to be (very) plausible.
>
> Minor Premise: The respondent is a bad person (ethically speaking), and therefore he is not credible.
>
> Conclusion: The respondent's argument should not be judged to be (very) plausible.

The critical questions for the direct *ad hominem* argument are the following (Walton, 1998a, p. 214).

> CQ1: Is the premise true (or well supported) that the respondent is a person of bad character?
>
> CQ2: Is the issue of character relevant in the dialog?
>
> CQ3: Is the conclusion of the argument the claim that the respondent's argument should be (absolutely) rejected, or is the conclusion merely the claim that the respondent's argument should be assigned a reduced plausibility?

The third critical question is important because *ad hominem* arguments, even when they are weak, have a "knockout" effect of getting the audience to discard an argument completely. But in many cases, a respondent's original argument may have been quite strong, and while the *ad hominem* attack raises legitimate questions about that argument, it by no means refutes it altogether.

6   Analysis of the cases based on embeddings

Both the PC case and the common currency case involve appeal to expert opinion. One key difference is that in the PC case the expert, Brent, is present to engage in dialog with Bill and Edith. In the common currency case, Dr. Smith's opinion is cited, but presumably Dr. Smith is not actually present to answer questions. Even though the same form of argument – appeal to expert opinion – was involved in the dialectical shift in both cases, the types of dialog were different. In the PC case, the shift was from a deliberation dialog to an information-seeking dialog. In the common currency case, the shift was from a critical discussion to an information seeking dialog. In this case, Helen and Roger are not deliberating on whether or not to adopt a common European currency. They are (we presume) not in a position to take action on this question themselves. They are having a kind of intellectual discussion on whether adopting a common European currency would be a good idea.

In the common currency case, the shift to the information-seeking dialog needed to help the original critical discussion on the European currency conversion issue. In judging whether European conversion to a single currency would be a good thing, examining past cases of currency conversion would, or at least could be relevant. But how exactly does the one dialog fit into the other? To see how, we need to turn to an analysis of argument alpha, which can be reconstructed as follows

Argument alpha
1. Conversion to a common currency improved the economies of all participating countries in all cases where it has been adopted.
2. Therefore conversion to a common currency will improve the economies of all countries in Europe who go ahead with conversion.
3. Improving the economies of all the participating countries is a good outcome.
4. Therefore, conversion to a common currency is a good idea.

Premise 1 was derived from expert opinion, gotten by consulting Dr. Smith. Premise 3, is a value-statement that would likely be generally accepted by both Helen and Roger. Premise 1 is a contentious claim that favors Roger's side. It was challenged by Helen in the dialog. Helen asked Roger why Dr. Smith holds this opinion, and then began a line of critical questioning that challenged the basis of the opinion. Possibly she could have carried this line of questioning even further by expanding the dialog. For example, she could have cited the opinion of a leading economist who disagrees with Dr. Smith's views on currency conversion. But even supposing that Helen accepts proposition 1, she could still argue that the inference from 1 to 2 does not follow. She could argue that conditions in all cases of

conversion in the past are different from the present situation in Europe. And that is in fact the line of argument she took.

The inference from 1 to 2 presumes an additional premise.

5.   All the other cases of currency conversion studied are similar to the present case of conversion to a common European currency.

It is also helpful to see that premise 1 was supported by appealing to the following prior premise put forward by Roger.

6.   Dr. Smith says so.

When Helen challenged the basis of Dr. Smith's reported opinion, Roger backed it up by citing Dr. Smith as an expert.

7.   Dr. Smith is an expert.

Putting this sequence of argumentation together, it can be represented as taking the form indicated by the argument diagram in figure 6.1.

Looking over the sequence of argumentation represented in figure 6.1, it is evident that the dialectical shift occurs at 2. Proposition 2 was the key premise in Roger's use of argumentation from consequences in argument alpha. But it was based on an appeal to expert opinion. The information-seeking dialog in which this appeal to expert opinion was questioned by Helen is represented in figure 6.1 by the part of the digraph above 2. So 2 is part of the information-seeking dialog, but it is also part of the critical discussion. It is the common point to both dialogs.

In the euthanasia case, the key form of argument is the slippery slope argument used by Wilma. This form of argument depends on a premise claiming that one step in a sequence leads to a next step, and so on in the sequence, until eventually some highly undesirable outcome occurs. Bruce's reply, citing the case of Holland as one where the sequence was stopped, apparently, provides a relevant kind of objection or critical questioning of the slippery slope argument. So, in a way similar to the argumentation in the currency conversion case, Bruce's statement about the case of Holland is a key proposition that links the information-seeking dialog together with the persuasion dialog. Why? It is because this proposition performs a dual role in the argumentation. It is a premise of Bruce's argument. But it is also a conclusion of a prior argument from expert opinion. It is at this point precisely that the dialectical shift occurs.
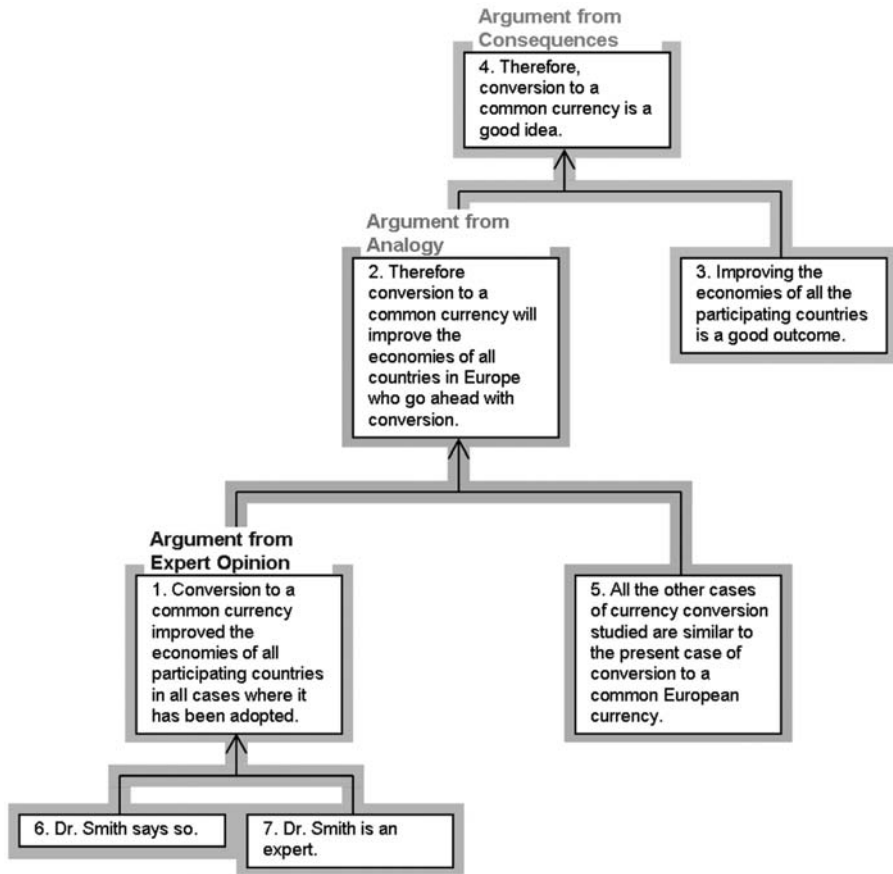
Figure 6.1 Argument diagram of the currency conversion case

The key thing to observe is that relevance can transfer, in some cases, over a dialecti-cal shift. For example, a proposition can be relevant in an information-seeking dia-log, because it represents an item of information sought in that dialog. For example, suppose it was found that a witness in a criminal case saw the defendant leaving the scene of a crime. But then suppose there is a shift to a critical discussion. Suppose that now the issue is whether the defendant committed that crime. Could that prop-osition, elicited in the information-seeking dialog, now be relevant in the critical discussion on the issue of whether the defendant committed that crime? The answer is clearly yes, in some cases. In other cases, however, a proposition elicited in an in-formation-seeking dialog might not be relevant, once there has been a shift to a persuasion dialog. For example, suppose we are discussing the issue of whether roses grow well in sandy soil, and you interject into the conversation, "The weather

in the English Channel today is calm but foggy". This information might have been relevant in response to some request for information in a previous conversation. But in the critical discussion about the growing of roses in sandy soil, it is not relevant. Once the shift to a different dialog occurs, the information is dialectically irrelevant. Or it is irrelevant as far as we know, unless further information or argumentation somehow connects channel weather with growing roses in sandy soil.

7   Analysis of the cases not based on embeddings

Like the cases based on embeddings, the shift in the bicycle case was abrupt, and can be specified exactly. But unlike the other cases, the line of argumentation being carried on in the first dialog does not continue through the shift. Instead, that line of argument was cut off, and an entirely different one began. The deliberation dialog in the interval is dialectically irrelevant to the critical discussion dialog on houses versus condominiums. In the other cases, the sequence of argumentation in the first dialog shared a common proposition with the sequence of argumentation in the second dialog. This structure was shown in the digraph in figure 6.1, in the common currency case. In the bicycle case, no such connecting digraph over the shift can be drawn. There was a discontinuity in the sequence of argumentation. It stopped at the shift, and then continued after the interval of the deliberation dialog was concluded.

The shift in the airline case led to a permanent blockage of the negotiation dialog. As the quarrel escalated, the personal attacks on each side got worse and worse. Finally Black called the union a "pack of morons" and White called Black a "slimeball". At this point, the quarrel has become so divisive and the language so derogatory that continuation of meaningful negotiations is blocked. Both sides are now too hostile to each other. Successful negotiation requires an empathy in which both sides understand not only their own interests, but also have some empathetic grasp of the wants and interests of the other party. There must also be a willingness to compromise. But to call the other party a "moron" or "slimeball" means that you no longer trust the other party, and you can no longer count on their ability to engage in rational argumentation. Hence the possibility of further negotiation is ruled out. This case illustrates the danger of introducing *ad hominem* argumentation that can be a problem in many types of dialog. The temptation is to reply to one *ad hominem* argument by using another in the "you are just as bad" pattern. As the shift to the quarrel escalates, there comes a point where there is no turning back. In this respect, the shift in the airlines case was worse than the shift in the bicycle case. At least in the bicycle case, the original dialog could be later resumed. In the airlines case, it could not be, and the outcome was a permanent blockage.

The main factor in the infomercial case is deception. The TV program was set up to look to the viewers like a consumer report. But then it shifted to a sales pitch. The purpose of setting up such a program is not altogether clear. It may be to try to convince the very gullible viewer that Pilot sunglasses are a consumer tested "best buy". But most of us know a commercial when we see one. More likely, the purpose was to just get the viewer interested initially, so he or she would sit through a half hour commercial. Also, the two types of dialog involved are complex.

The first dialog in the program can perhaps be classified as a kind of information-seeking dialog that reports on the results of tests made on the various sunglasses selected for testing. The second type of dialog is a sales pitch where the goal is to get the viewers to buy a product. A sales pitch is a kind of advocacy dialog. The respondent knows that the proponent is advocating a product she is selling. The respondent does not expect a two-sided, balanced and objective dialog. But in a consumer's report type of dialog, the respondent does expect balance and objectivity. If balance and objectivity are compromised, by pushing one product and trying to promote it, the credibility of the report would be damaged. From a viewpoint of the critical analysis and evaluation of argumentation, the shift is significant in the infomercial case, because it goes from a two-sided dialog in which objectivity is expected to an advocacy dialog that is one-sided.

In the dialog in the Mexican war case, Harry made a claim that Mexico provoked the United States. Louise then expressed doubt about this claim, and asked Harry to prove it. In a critical discussion of the sort Louise and Harry are taking part in, there is a burden of proof on any party who makes a claim. If the claimant is asked to prove the claim, or at least provide evidence for it, then the claimant must either offer such evidence or he must retract the claim (Walton and Krabbe, 1995). In this case, Harry failed to fulfill this requirement. Indeed, at the next move, where he was supposed to fulfill this requirement, he shifted to an altogether different type of dialog which went into practical questions of the consequences of holding a view. In the language of fallacies, it could be said that Harry's move in the dialog was irrelevant. He tried to avoid fulfilling the requirements of burden of proof by switching to a quite different issue. In this case, dialectical irrelevance occurs precisely because of the shift from one type of dialog to another. The shift is masked because the accusation made against Louise is emotional and distracting. She is accused of being unpatriotic, and even of doing something very bad in the form of "giving comfort to our enemies". In this case, the accusation is so implausible that it is more funny than serious. But it is not hard to see that in some cases such a reply could be highly provocative and distracting.

Argumentation from consequences is often quite a reasonable argument, but in this case its use is fallacious. Underlying the committing of the fallacy of irrelevance in the case is the dialectical shift. The deliberation dialog, citing bad conse-

quences of holding a view, is not a good reason Harry can use to support his line of argumentation in the dialog. Louise's asking for proof is a relevant question. Harry needs to either give proof or give up his claim, to carry on the persuasion dialog properly. Instead, there was a shift. There was no embedding of the second type of dialog into the first. Harry used argumentation from consequences in the form required by the argumentation scheme. He cited bad consequences as an outcome of what Louise was doing – holding a certain view. But in so doing he failed to properly answer Louise's question, and thereby created a rupture in the persuasion dialog. From the normative standard of persuasion dialog, what he did was an improper move. Harry broke a rule of critical discussion, by failing to make the required kind of move in response to Louise's question. His reply was not constructive in moving the persuasion dialog forward. Hence there was no embedding. In this case, the shift occurred right in the use of argumentation from consequences when Harry made his reply to Louise's question asking for proof.

8   Fitting dialogs together at global and local levels

There are two levels that must fit together for an embedding to occur. One is the global level. The goal of the second type of dialog must be compatible with the goal of the first type of dialog. For example, suppose the first type of dialog is a deliberation on how to take some practical course of action in a given situation. Information is needed if the deliberation is to be successful. For we typically need to know what the situation is, how a task can be carried out, and what are the likely consequences of carrying it out or not. Without such information, the deliberation could be much worse. Generally then, the goal of information-seeking dialog tends to be quite compatible with the goal of deliberation. If it is the right sort of information, such information can be extremely useful in a deliberation. One particular type of information-seeking dialog, the expert consultation dialog, can also be extremely useful to support the goal of either a critical discussion type of dialog or a deliberation. The cases studied above illustrate how this fitting of goals can work when there is an embedding. The other level to be considered is the local level. In studying a case, it is crucial to examine the local area in the dialog right around where the shift occurred. This local area can be a single move, as in the Mexican war case. Or it can be a sequence of moves, as in the airline case.

The airline case illustrates the kind of case in which there is a failure to fit at the global level. Generally, the shift from a critical discussion to a quarrel is illicit, because quarreling is simply not an efficient way to conduct a critical discussion. In a critical discussion, the aim is to resolve the original conflict of opinions by having each side present rational evidence to support its claims. Mounting a per-

sonal attack against the other side is, generally speaking, not a good way to meet this burden of proof. If we are having a critical discussion on the issue of physician-assisted suicide, for example, my attacking you by claiming you have a bad character for honesty contributes nothing to supporting or refuting the thesis that physician-assisted suicide is good or bad. It could possibly be relevant, depending on the local particulars of the case. But globally, it is not a good fit. It doesn't advance the critical discussion, by itself, and seems to be irrelevant in that context.

In short then, there are two levels at which a shift needs to be evaluated – the global level of dialog goals, and the local level of specific moves in the dialog around the shift in the given case. At the local level, it needs to asked what question an argument was an answer to, and other matters of that sort. One factor that comes sharply into play at the local level in many cases is the argumentation scheme. The scheme is a normative structure that can be applied at the local level to determine many crucial factors needed to evaluate a shift. The scheme determines what premises need to be supported to make the argument carry weight in the dialog. So, for example, if the argument is an appeal to expert opinion, one premise reports a claim made by the expert, ostensibly reporting some factual information. In such a case, information-seeking dialog fits in very naturally with the prior dialog, as illustrated by the common currency case. What makes the embedding work is the local use of the form of argument called appeal to expert opinion. This argumentation scheme is the local structure that enables the one dialog to be fitted constructively into the other. In such a case, the fit is local as well as global.

The shift from persuasion dialog to deliberation dialog is generally not a good fit, although there is no reason to think it is always illicit. In two of the cases studied here, the shift is not based on an underlying embedding. In the bicycle case, there is no local fit in the argumentation around the shift. The shift to the deliberation interval does nothing at all to improve the critical discussion on the issue of house versus condominium. There is no line of argumentation at the local level continuing through the shift. Hence there is no embedding in this case. In the Mexican war case, it looks on the surface like the line of argumentation continues at the local level over the shift. Harry uses argumentation from consequences, a form of argument that is in principle relevant in the critical discussion on who was in the right in the Mexican war. But his use of this argument is problematic precisely because of the shift triggered by the way he used the argument.

There is a shift to a different level where the consequences of Louise's holding her viewpoint were brought into consideration.

In the infomercial case, the shift is indicated by an imperative sentence, "Get a pair right away." The goal of the sales pitch is to get action. But the consumer's report type of dialog is reporting on tests of the product. In that type of dialog, an imperative sentence, or a sentence that directs the listener to a specific course of

action, like going out and buying a specific brand of sunglasses, is out of place. So when Zelda says, "Get a pair (of Pilot sunglasses) right away.", we realize that a dialectical shift has taken place. But notice that there are several prior indicators in the dialog that lead into the shift. Zelda says that Pilot sunglasses "really look great". Then Dan says, "They are priced very reasonably, considering their high quality." These statements are still relevant in the consumer's report type of dialog. A consumer's report would consider the prices of products, and might consider how the product looks. But then, at the next move, Zelda says, "You can get them for $12.98 at all Topline Store outlets today." This statement, which tells the viewer where to buy a specific product, and what it costs, begins to sound suspiciously like an ad. The reason is that it gives information that shows the viewer where to purchase a specific product that is being enthusiastically endorsed at this point. It begins to sound like a commercial. The next step is taken when Dan even gives a toll-free number. All these prior steps gradually lead up to the shift at the final move in the dialog. Gradually there is more of a preparation for a specific recommendation for action that is being led up to. So when the shift actually occurs, it is fairly clear what is going on. The transition from more of a descriptive type of language to an imperative type of language is an indicator of the dialectical shift.

In the Mexican war case, as noted above, the shift is quick, taking place at one move. Also, when Harry makes the shift, he is still talking about the issue of whether the United States had justice on its side. So even after the shift, what Harry says is topically relevant to the issue of the dialog before the shift. In this case then, the shift is subtle. It is harder to detect than the shift in the sunglasses case. We know that there has been a dialectical shift in the Mexican war case, once we realize what has happened. But what specific indicators are there that one could look for? The answer is that there is a similarity with the sunglasses case. In both cases there is a shift to an imperative language that directs the respondent to a course of action. In the Mexican war case, Harry argued at the last move in the dialog that Louise's questioning his view will "give comfort to our enemies by promoting the cause of defeatism." What Harry is claiming is that if Louise continues to question his view that the United States had justice on its side, bad consequences will follow. He is using argument from consequences. What is the conclusion of this argument from consequences? It is the statement that Louise should stop questioning his view. This statement is a practical directive. It directs Louise on prudential grounds to stop questioning a view, because of the claimed bad consequences of continuing to question. Hence in this case, it is also possible to detect a shift to an imperative type of language characteristic of practical discourse. Harry is arguing that Louise should give up her questioning about which side had justice on its side in the Mexican war. But at the shift point, he begins to do so on practical grounds of the consequences of her holding this view. He is no longer talking

about whether his view is justified by the facts about the Mexican war of 1848 or not. He is talking about Louise's actions, and what she should stop doing, on grounds of the consequences of not stopping.

In both cases, an indicator of the shift is a change in the language to an imperative kind of language that directs one party to a specific action. The shift in both cases is to practical reasoning of a kind in which the conclusion is a directive to action. The argument becomes a recommendation by the proponent to the respondent to take some course of action alleged by the proponent to be prudent for the respondent. Recommendations for action are generally not relevant in an information-seeking dialog or a critical discussion (subject to exceptions). In some formal models of persuasion dialog, recommendations would not even be allowed as permissible locutions by the locution rules. For example, in the formal model of persuasion dialog PPD_0 the permitted locutions include assertions and statements, but do not include imperatives or recommendations for action (Walton and Krabbe, 1995, p. 149). In a persuasion dialog, of at least certain kinds, the making of a recommendation for action is irrelevant. So as soon as one party begins to make recommendations that the other party take some indicated form of action, there is an indicator of a dialectical shift. The purpose of a critical discussion is to resolve a conflict of opinions. The issue is whether a particular proposition is true or false. If there is a shift to practical reasoning about actions, that is an indicator of a shift to a different type of dialog. Such an indicator is not absolute, as some kinds of critical discussions can be about actions in some cases. But typically, and in many cases, it is a sign or indicator that may be evidence of a shift.

The problem in both the sunglasses case and the Mexican war case is one of relevance. The fallacy in both cases could be diagnosed as one of irrelevance. The Mexican war case is particularly interesting in this respect, because, as noted above, the wording in Harry's last move in the dialog is topically relevant. How then can we say that Harry's use of argument from consequences is a failure of relevance? The first step is to recognize that dialectical relevance is not the same thing as topical relevance. Harry's last move is topically relevant. The reason his move is dialectically irrelevant is the dialectical shift from the critical discussion to the practical argumentation from consequences – a kind of argumentation that would be relevant in a deliberation but that is not relevant to the critical discussion that Harry and Louise were previously engaged in. The basis of the dialectical irrelevance is the failure of the practical discourse to be properly embedded in the prior critical discussion on whether the United Sates had justice on its side in the Mexican war of 1848. This type of dialectical shift is common enough to be generally interesting to those studying fallacies. It can happen in comparable cases where, for example, the literary merits of a film or novel are being discussed. The question may be put as "Was it a good film?" meaning was it good from an aes-

thetic perspective? But then one party to the discussion claimed that the film incited violence, by suggesting a type of criminal activity to teenagers. The shift here is to a kind of practical discourse based on argument from consequences. A failure to detect the shift can result in the fallacy typically called "irrelevant conclusion" in logic textbooks (Krabbe, 1992). Because the film had bad consequences, it might be an error to leap to the conclusion that the film was aesthetically bad. Examples of arguments of this sort can be quite difficult to analyze and evaluate. Each case needs to be studied on its individual merits. But the broader significance of this kind of case is worth pointing out.

9   Metadialogs

A metadialog (Krabbe, 2003, p. 83) is a secondary dialog about a ground level dialog that was already going on, so that the metadialog was produced by a dialectical shift from the ground level dialog. For example, a persuasion dialog might be underway, but then there might be disagreement between the participants about the correctness of some moves that were made in the persuasion dialog. To solve the problem, so that the persuasion dialog can proceed, the participants may move to a metadialog to have a secondary dialog on whether the move in the first dialog can be judged to be correct or not by some criteria (Hamblin, 1970; Krabbe 2003). Some shifts are easy cases that can be managed without undue effort by simple rules that can be more or less automatically applied to the case. Others are hard cases that require systematic intervention, and cannot be fairly adjudicated without the intervention of a third party at a metadialog level.

On the view of Wooldridge, McBurney and Parsons (2005), argumentation in dialog is inherently meta-logical, meaning that it does not just involve the putting forward of arguments at one level of discourse, but also the putting forward of arguments about these arguments at a higher level. As an example they consider an argument that takes place between advocates in a trial. At the meta-argument level, arguments can be made by the judge about these arguments that took place at the argument level. Wooldridge, McBurney and Parsons (2005, p. 7), note that meta-logical systems have been widely studied in the past four decades in artificial intelligence and logic. Prakken (2001) constructed a formal system to show how shifts in a burden of proof work in legal reasoning of the most common sort, classifying such problems as irreducibly procedural (p. 253) that can only be solved by turning to metadialog considerations. A common example in legal argumentation is how a burden of proof can shift back and forth in a trial, requiring a decision on burden of proof by the judge. Understanding such a shift is a central problem for argumentation theory as a whole, one best approached through the study of metadialogs.

Krabbe (2003, p. 83) formulated the demarcation problem of deciding in a given case of argumentation where there has been a shift to metadialog which moves belong to the ground level and which ones belong to the metadialog level. As we saw in section 1, a burden of proof impasse can occur in cases involving argument from ignorance. Suppose one side says "You prove it!" and the other side replies "You disprove it". If trying to resolve such a dispute within a persuasion dialog proves impossible, a shift to a different type of dialog may help. The diagram in figure 6.2 shows how a shift from an original dialog D1, at a ground level, can lead to a metadialog that then begins at a next level where a secondary dialog D2 begins.
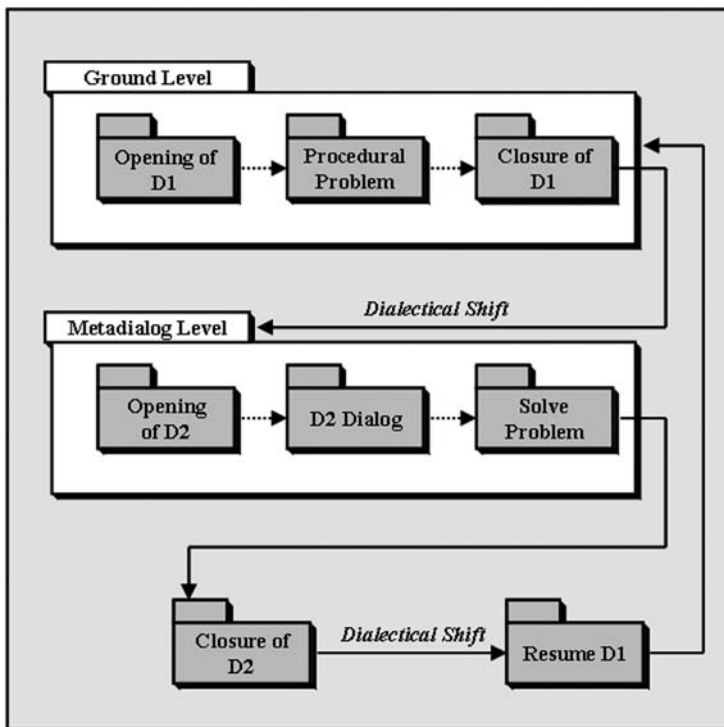
Figure 6.2 Shift from a ground level dialog to a metadialog

Solving the demarcation problem requires the application of dialog identifiers to the two types of dialog involved, recognizing their stages, and judging how each of the points in figure 6.2 applies to the sequence of argumentation. Krabbe (2003, p. 83) identified two other problems. The problem of infinite regress arises when a disagreement about ground level rules in a dialog leads to a metadialog, which in turn leads to a secondary metadialog about a procedural problem arising in the

initial one. This problem could be an infinite regress, or circular argumentation that hops back and forth from the one level to another and back again.

10   Solving the embedding problem

Grasping that a dialectical shift has occurred in discourse is an intuitive judgment we make as native speakers of a language. For example, when a consumer's report type of discourse shifts to a commercial, we can see a shift has occurred. We know a commercial when we see it, because we are very familiar with this particular type of speech event. But how can we teach someone to become more aware of such shifts? And what evidence from the text of discourse of a given case can be used to back up or refute the claim that a shift has taken place? In many cases of everyday conversational argumentation, it is difficult or even impossible to determine whether a dialectical shift has occurred. For in many such cases, there is no clear prior agreement or understanding of what kind of conversation the two parties are engaged in. Conversations in real cases are often undirected, wandering and un-finished. In some cases, people even try to deceive each other about the supposed aims or direction of a conversation. Even so, when it comes to evaluating cases of argumentation, especially one associated with the committing of fallacies, it can be very useful to give a conditional evaluation of an argument. A conditional evalua-tion takes the following form: if the speaker was supposedly making a contribu-tion to a particular type of dialog that can be identified, then his argument falls short of the proper requirements for that type of dialog in this particular respect.

Solving the problem of embeddings requires a fitting together of several fac-tors in a given case. One factor is the type of dialog involved in both the first and the subsequent dialog sequences, before and after the shift. A global judgment needs to be made on whether the shift is an embedding or not, based on the goals of the two dialogs and the order in which they occur. Another factor is the pin-pointing of the shift at the local level. As observed above, some shifts are gradual and others take place at a specific move in the dialog sequence. In some cases, a gradual movement toward the shift could be detected, but then, at a particular point, the shift occurred. Certain types of shifts, like many of the ones cited in the cases above, are very common and typical. By studying these common cases, we can learn to recognize more subtle shifts when they occur in other cases. Another factor concerns the relationships of the local moves in a dialog. For example, in the Mexican war case, at her last move in the dialog Louise asked Harry "How can you prove that claim?" Therefore, according to the rules for a critical discussion (van Eemeren and Grootendorst, 1992), Harry, at his next move, must give an argu-

ment furnishing evidence to support his claim, or else he must retract the claim. Does his reply fulfill such a requirement? No, it does not. But it is a subtle matter to pinpoint precisely why not. The reason is to be found in the dialectical shift. But the subtle aspect is that Harry is giving an argument – an argument from consequences. And so it may appear on the surface that he has met the appropriate requirement for the critical discussion. But in reality, he has not. He has shifted to a different type of dialog. His response is relevant in that type of dialog – a deliberation. But it is not relevant in relation to the original type of dialog that Harry and Louise were supposedly engaged in – the critical discussion about whether the Untied Sates was in the right in the Mexican war of 1848. The fallacy of irrelevance is best evaluated by seeing why there is a failure of embedding in the case.

What determines whether an embedding exists is whether the argumentation that occurs after the shift is useful to contribute to the goal of the dialog that was taking place before the shift. Another factor that can be important in determining whether an argument in a secondary dialog is useful for the primary (prior) dialog is whether the subject-matter of the argumentation in the secondary dialog fits into that of the first. For example, suppose that in the critical discussion on euthanasia Bruce starts reading an article on the building of dikes in Holland. Both the subject of physician-assisted euthanasia in Holland and the building of dikes in Holland have a topic in common – Holland. But what use is information on the building of dikes in Holland in relation to the critical discussion on euthanasia that was going on before between Bruce and Wilma? None at all, it would appear. Wilma could, and should ask at this point, "Is that relevant?" The shift, unlike the shift in the euthanasia case, does not appear to be an embedding, unless Bruce can somehow make a connection showing the dike information is useful in the discussion on euthanasia. In the absence of such a connection being made, the conditional evaluation should be that the information introduced by Bruce is irrelevant.

In short, the various factors that make for an embedding are multiple. Not only must the goal of the second dialog fit with and support the goal of the first. Many other factors must fit in over the sequence of argumentation connecting the two dialogs. The right grammatical parsing of the moves must mesh. For example, a shift from assertive sentences to imperative sentences may indicate a shift. Also, the moves must match each other at the local level. For example, asking for proof of a claim requires that an argument be given at the next move, or some other kind of reply that is appropriate as a next move. Whether or not the next move matches its prior move is determined by the dialog rules (Krabbe, 1992). Also, the digraph of the sequence of argumentation over the shift must fit together showing a functional fit between the segments on either side of the shift. Another important factor is the argumentation scheme. The cases above show how a shift often occurs within the use of a particular type of argument, like an appeal to expert opinion.

The argumentation scheme for appeal to expert opinion indicates that various kinds of consideration are relevant, because they are requirements that are parts of this type of argument, as indicated by the argumentation scheme. All these factors must fit together in the right way in a case, so that the argumentation after the shift is useful to contribute to the goal of dialog before the shift. These factors can be summarized as follows.

1. The respective goals of the two dialogs, and how the goals fit together.
2. Grammatical parsings of the moves adjacent to the shift
3. Matching of a pair of moves at the local level where the shift occurred.
4. Fitting together of the sequence of argumentation over the shift.
5. How the local moves fit into the argumentation scheme for the argument used at the shift.

To judge whether a shift is based on an embedding, evidence relating to each of these five different factors can be important. Much depends on what types of moves were being made right around the local area of the shift, or right at the shift. If a particular type of argument – like a slippery slope argument, or appeal to expert opinion, or argument from consequences – was put forward at or right near the shift, then the argumentation scheme for that type of argument will be vital in determining whether the shift is an embedding.

The solution to the problem of embeddings proposed has two parts. The first part concerns the relatively clear kinds of cases. The second part concerns the more problematic kinds of cases. The clear cases are of various sorts. In some cases, the participants explicitly agree to change to a different type of dialog. In other cases, although no explicit verbal agreement was made, nevertheless it is quite clear to participants and onlookers alike that a definite shift has occurred. An example of this sort is the bicycle case. In these clear cases, there is comparatively little difficulty in dealing with the shift. The shift can be marked at the appropriate place in the sequence of dialog. The system of (Reed, 1998), using notation to mark each type of dialog, is what is required for this purpose. In the formal model of the dialog, each shift will be visibly marked. And then anyone examining the dialog can examine the evidence of the case and judge, on this evidence, whether the shift is an embedding or not. Consider Reed's case of picture-hanging (Reed, 1998, p. 249) cited earlier. Here is a simplified version of how Reed models the sequence of dialog. White proposes that he and Black hang a picture. During the deliberation, Black accepts the proposal for hanging the picture. White can hang the picture, because he has the tools. All that he lacks is a nail that can be supplied by Black. White and Black then shift to a negotiation dialog. White proposes that Back supply the nail. Black then accepts the proposal.

> White: *propose* (*deliberate* (*can*, *White*, *hang-picture*))
>
> Black: *accept* (*deliberate* (*can*, *White*, *hang-picture*))
>
> White: *propose* (*have* (*White*, *nail)* and (*can*, *White*, *hang-picture*))
>
> Black: *accept* (*have* (*White*, *nail*) and (*can*, *White*, *hang-picture*))

In this brief dialog, there is a shift from deliberation to negotiation. What is interesting to note especially is the use of the propose-accept notation. In the framework of van Eemeren and Grootendorst (1992) and Walton (1998), the type of dialog is determined at the opening stage in a kind of meta-discussion in which the participants agree on what type of dialog they will engage in. Reed's notation brings this meta-discussion into the local level. At a specific move, participants can decide to move to a different type of dialog. And there is a clear representation of any dialectical shift right at the local level.

This approach doesn't, of course, solve all the problems of shifts encountered in the case studies above. Participants might not agree to shift to an embedded type of dialog when one is needed. Or a participant might propose shifting to a non-embedded type of dialog, and the other participant might agree to it. The problem still remains of deciding whether a proposed shift is an embedding or not. Another kind of problem is that a participant may not propose a shift, but then make the shift anyway to a different type of dialog, without announcing it, or securing agreement from the other participant. In this sort of case, the five factors cited above must come into play. In a formal system of dialog, presumably a shift to a second type of dialog would violate the rules for the first type of dialog. Thus through such a failure the shift would be detected.

In order to solve the problem posed by the fallacy of bargaining (Wells and Reed, 2006) proposed two new formal dialectical systems, one of which represents persuasion dialog while the other represents negotiation dialog. The components for each system include an identifier for the system, a structure for taking turns, identifiers for the participants and the setting up of commitment stores. Most importantly, each type of dialog includes rules that state conditions for entering a new dialog, rules that govern transitions from one dialog to another, and rules for specifying when a dialog should terminate. Using this formal structure, Wells and Reed show how the fallacy of bargaining involves an illicit shift from a persuasion dialog to a negotiation dialog. To do this, they state formal conditions under which one dialog can be embedded in another, and also show how these conditions are violated when there has been an illicit dialectical shift from one dialog to another.

The second part of the solution relates to the more problematic kinds of cases in our case studies. These are real cases of natural language dialog, where not all shifts are clearly marked or agreed to. In some of these cases, the participants made

no explicit agreement to shift to a different type of dialog, and while there is some evidence of such a shift, the evidence is not conclusive. A case in point is the Mexican war case. The use of practical language concerning actions strongly suggests a shift from a critical discussion to some more practical type of discourse like deliberation. Another problematic case is the sunglasses case. There seems to be a shift from a consumer's report type of dialog to a sales pitch. But neither of these types of dialog is clearly defined or categorized by the system of classification in Table 1.2. Even so, the shift does not seem to be an embedding. But to back up this analysis by showing how the one dialog fails to fit into the other is problematic. The best that can be done to deal with such a case is to take all five kinds of evidence into account in evaluating, relative to the given particulars of the text and context of discourse in the case, whether a shift represents an embedding. Each case should be judged on the basis of the evidence. And of course, in some cases, it will be hard to tell whether a shift has really occurred or not, and whether the shift really is an embedding or not. However, in the cases studied above, there was sufficient evidence available from the text and context of discourse for an analyst to offer, if not a decisive, at least an informative evaluation. The evidence from the case can be assembled and evaluated. If there is ambiguity, the analysis can point out the ambiguity. The awareness that the case can be evaluated two different ways is itself often very illuminating. Formal systems of dialog are useful in such cases, in order to present a normative model of how argumentation should ideally go. But of course, any realistic case in a natural language text of discourse may only meet some, and not all, of the requirements of the formal model. In such cases, a conditional evaluation may be the best that can be accomplished.

Even a conditional evaluation can be quite useful in alerting a critic to serious problems. Such cases of everyday argumentation requiring conditional evaluation are fairly common. What is worth noting is that many of them, as illustrated by the cases studied above, contain types of argument associated with traditional informal fallacies. When what are taken to be fallacies occur in realistic cases of natural language dialog, it may be necessary to make numerous dialectical assumptions to get a workable interpretation of the text of discourse. But of course, it is just in such cases that the ability to pinpoint shifts and embeddings is most useful for applied logic. What is shown, even by the relatively simple cases studied above, is that the study of shifts and embeddings throws useful illumination on problems of analysis of arguments associated with the informal fallacies. The most promising connection may be the relationship between the study of shifts and embeddings with problems of relevance. The most significant indicator of irrelevance (of the kind associated with fallacies of irrelevance) is the occurrence in a given case of a shift not based on an embedding. Among its other uses, the study of embeddings opens up new avenues for the identification, analysis and evaluation of fallacies of irrelevance.

Criticizing a natural language argument

Modern argumentation theory is based on the assumption that arguments as presented in a natural language text of discourse can be analyzed and evaluated from a critical point of view. This assumption is, in turn, based on the assumption that a natural language text of discourse can be interpreted by the would-be critic as expressing some clearly attributable argument or set of assertions. But is this secondary assumption justified? Or, at least, are there some methods in the process of development that can make this assumption worth adopting? There are problems in attributing arguments to an arguer, based on a given text of discourse. One of these problems is the so-called straw man fallacy. The tactic used in this fallacy is to wrongly attribute a more extreme position to an arguer than the evidence in the text of discourse merits, and then use that misrepresentation to try to refute the arguer. It is often said that an argument needs to be based on a charitable or friendly interpretation before it can be properly criticized, as opposed to merely being attacked in a quarrelsome way (Johnson, 1981). The problem is how this task can be precisely carried out using some objective methods that will guarantee that a criticism is not based on an unfair interpretation, or even one that commits the straw man fallacy.

In this chapter, it is argued that there is a dialectical method of textual interpretation useful for critical analysis and evaluation of argumentation. The method is based on dialog theory, and uses a process of logical reasoning to construct a hypothesis from a textual data base that functions as a given, reproducible body of evidence. The method is also based on plan recognition technologies developed in artificial intelligence, and new tools for argument diagramming. These tools can be used to show how different interpretations of an argument can be constructed, and how one interpretation can be shown to be more reasonable, or more in line with the commitments the arguer has expressed in the given text of discourse, than another interpretation. Thus the method produces results that are verifiable and objective, or so it is argued in this chapter. The method is based on the structures of dialog that have been developed in the previous chapters of this book, on plan recognition technology, on argument diagramming tools, and on abductive inference, or inference to the best explanation, as it is often called. The method is

also built on some maxims and principles of textual interpretation derived from traditions of religious hermeneutics and statutory interpretation in law.

Studying the problems of fair criticism and argument analysis requires beginning with some easy cases that can be used to build objective methods and tools that can then later be applied to hard cases. There is no space in this chapter for lengthy consideration of hard cases. To illustrate how the methods work we began with a series of easy cases, and then move on to articulating the general principles that need to be used to approach harder cases. The aim is not to build new argument technologies in addition to the ones already mentioned in this book and in the first part of chapter 7, but to build a general theory of dialog to be used as a foundation for the building of such systems and for developing computational devices in the future. The purpose is to extend the theory of dialog developed in the previous chapters to the task of fairly criticizing an argument found in a natural language text of discourse.

1   Explanation, clarification and interpretation

Much of the problem of applying logic to actual cases of arguments, as Alfred Sidgwick (1910) clearly showed, is prior to the problem of judging whether the given argument is weak or strong. The prior problem is answering the question, "What is the argument?" For typical of evaluating arguments of the kind associated with informal fallacies is dealing with ambiguity, vagueness, nonexplicit premises, and other matters of how an argument is presented or put forward in a text of discourse. It is conceded even in traditional formal logic that many arguments are enthymemes, or arguments with premises or conclusions that have not been explicitly stated. Very often it is not hard to determine whether or not an argument is valid or invalid, once it has been clearly specified what its premises and conclusions are. The more difficult part is to get agreement on what the argument should precisely be taken to be, given the text of discourse that allegedly contains or expresses it.

Trying to identify, analyze and evaluate an argument used in a given case is based on what is called the interpretation of the text of discourse. Interpretation is not a trivial task. It involves making many assumptions or hypotheses that cannot be proved with certainty, or even with high probability. However, evidence from the text and context of discourse in a given case does bear on such hypotheses. In principle therefore, it would seem, such hypotheses can be verified or falsified. At least they can be supported by the evidence, or refuted by the evidence. But much depends on what we take 'interpretation' to mean. Interpretation should not be taken to mean determining what a real arguer actually meant to say or believes.

That is a psychological task. Interpretation has the lesser goal of merely determining what the arguer should be taken to mean, judging from what she said, and from what is known from the context of the conversation that her argument was part of. The goal of interpretation should be to judge what the arguer's commitments are, as can be implied from what she has gone "on record" as saying, given the actual wording of the text of discourse in the given case. In other words, interpretation tries to determine commitment. And commitment is different from belief. Commitment is acceptance, as opposed to belief. Belief is a private psychological matter which can change at any time. Commitment is fixed, though not inalterably, by what one has said, or written down, in a given case (Walton and Krabbe, 1995). Commitment leaves tracks in argumentation, speech or writing, and these tracks yield the evidence needed to identify, analyze and evaluate the argumentation.

Commitment is partly a subjective notion, because it is related to point of view. A point of view (viewpoint, standpoint) has two components – a proposition and an attitude (van Eemeren and Grootendorst, 1984). An attitude can be pro, contra, or neither. If neither, it is said to represent the so-called zero or neutral point of view (van Eemeren and Grootendorst, 1992). In the task of interpreting, analyzing and evaluating an argument, two points of view are involved. One is the point of view of the critic, or the person who is trying to identify, analyze and evaluate the given argument. It is very important to distinguish between a critical point of view and an opposed, or attacking point of view. The critic should not adopt a contra point of view, of being against the given argument and trying to attack it. Unfortunately, criticism is often taken this way. The critic should take a balanced point of view. On the one hand, the critic does need to find the weak points in the given argument. He needs to do this for constructive purposes of trying to open up possibilities of improving and strengthening the argument, based on criticisms of it. But the critic also needs the important quality of empathy. The critic must try to appreciate the strong points of the argument by entering into the point of view of the arguer. The second point of view involved is that of the arguer herself. The arguer has a point of view, as expressed by the argumentation she has put forth. To sum up then, the task of identifying, analyzing and evaluating an argument involves the juxtaposition of two points of view. What needs to be determined and articulated is the arguer's point of view. To do this, the critic must adopt what is called a critical point of view.

It is very hard for some beginners to logic to grasp the meaning and requirement of a critical point of view. At first, many people think that you must be either for or against an argument. Of course, this is true. As a human being with personal preferences and special interests, you probably will be for or against it. But to be a critic, as required for purposes of logically evaluating an argument, you have

to suspend or bracket these feelings, to the extent you can. You have to put yourself in the viewpoint of the arguer, in a hypothetical act of empathy, to get some grasp of what she is "trying to prove". Even if you don't agree with a viewpoint, you can still have some grasp of how it hangs together and where it would plausibly go as a line of argument. What has to be developed is a notion of constructive criticism. Criticism should not be seen, in line with this notion, as an attempt to attack an arguer, or to refute or destroy his argument. It should be seen as a kind of critical questioning that can probe into the weak points, or unclear points of an argument, and thereby reveal specific respects in which the argument can be strengthened and improved. In some cases, constructive criticism will have the effect of refuting an argument. But that could be a constructive outcome, if the existing argument can be modified, or some better line of argument chosen, as an alternative to the existing argument. So while there is an adversarial element to criticism, there should also be a collaborative element. For the collaborative element to work properly in constructive criticism, both parties, the critic and the arguer criticized, need to be open to recognizing deficiencies in their arguments, and to recognizing merit in opposed arguments. Despite all these sentiments however, accepting criticism is not easy for most of us. It is a delicate art to balance advocacy of one's own argument against the need to take legitimate criticism of it into account.

It is often taken for granted that interpreting a natural language text of discourse is purely subjective, and that therefore there is no objective evidence that could be used to prove or disprove a particular interpretation is right or wrong. In fact, in practice we are quite skilled in interpreting the discourse of a speech partner, and such a skill is fundamentally important to successful communication, and probably even to our survival. We not only do it very well (although with impressive and significant failures on some occasions), but we can also find useful methods for doing it systematically for purposes of criticizing arguments. A lot of the problem with our failure to discover such methods in the past has been our failure to realize that arguments can be put forward for different purposes in different conversational settings. Determining what an argument is, or should be taken to be, may depend significantly not only on what was said, but on the purpose of saying it within a conversational setting. Such conversational settings, or types of dialog, do have structures, goals and rules. Knowledge of such a conversational setting helps a critic to determine what the argument is, and what its non-explicit parts should be taken to be.

The following factors are important in interpreting argumentation in natural language discourse from a critical point of view.
1. The finding and resolving of ambiguity in key terms.
2. The clarification of definitions of key terms.
3. Judging whether something is an argument versus, say, an explanation

4. Determining the premises and conclusion of an argument.
5. The reconstruction of chains of argumentation, using an argument diagram.
6. The deletion of matter that is of secondary importance.
7. The insertion of nonexplicit parts of an incomplete argument, like missing premises.
8. The contextual clarification of the type of dialog an argument is part of.
9. The determination of the ultimate conclusion the argument is leading to.
10. The fixing of the extent and content of the text of discourse of the case.

The assumption is that if we are presented with a natural language text of discourse in a given case, the limits of it can be determined. Is it a letter to the editor, or is it a long article, or even a whole book? To begin to analyze any case, the limits of the case need to be established. This given text of discourse, along with what is known of its context, provide the evidence against which all interpretations of the argumentation in the case must be tested. Assumptions can be made about what was plausibly, or most likely meant in a given case. But all such assumptions need to be backed up by evidence from the actual wording of how the argument was expressed in the text of the case. In some cases, there is evidence both for and against a given interpretation. But that should be expected. People often retract their commitments in argumentation in real cases. They sometimes even intentionally try to deceive an audience by exploiting ambiguity or confusion. So an interpretation should not be seen as absolute. However, it must always be seen as being testable by the evidence furnished by the text of discourse in the given case.

At this point a careful distinction needs to be drawn between clarification and explanation. According to the new dialectical theory of explanation (Walton, 2004; 2007), an explanation is a speech act of a special kind that takes place in a dialog in which one party has the goal of transferring understanding to another party who lacks understanding about some event that appears puzzling to the party who has asked a question about it. On the new dialectical theory, there are three kinds of conditions that define what an explanation is: dialog conditions, understanding conditions and success conditions. According to the dialog conditions, two parties are engaged in some type of dialog that has collaborative rules, and in this dialog one party, whom we will call the explainee, has asked a specific question about something that appears anomalous to him. According to the understanding conditions, the explainer is presumed to have some kind of special understanding about the subject matter of the question that the explainee lacks. According to the success conditions, the explainer is supposed to supply a chain of reasoning that enables the explainee to come to understand the anomaly he previously claimed that he failed to understand.

The speech act of clarification is similar to the speech act of explanation in many respects. Indeed, a clarification can technically be seen as a species of explanation, and both speech acts need to be defined by comparable conditions. What makes clarification different from explanation is that clarification always takes place in a situation where one party in a dialog has made some kind of move involving assertions that are part of a text of discourse that this party is responsible for, and the speech act of making a request for clarification put forward by the other party always refers back to some part of that text of discourse put forward by the first party. In other words, a clarification is a special kind of explanation requested to a party in a dialog to explain something she has previously said in that dialog. The other distinguishing feature of clarification that differentiates it from explanation as a speech act is that clarification needs to be seen as having the structure of a metadialog, meaning a dialog about another dialog (Krabbe, 2003). A clarification always involves a shift from a given dialog that is taking place to a higher level in which a secondary dialog starts to occur about something that transpired in the first dialog.

Finally we come to the concept of an interpretation. The kind of interpretation we have in mind is one that is appropriate for critical argumentation. In critical argumentation, two parties are involved. One party has put forward an argument, and the second party has taken on the role of critic of that argument, not meaning 'critic' in the negative sense of someone who is only attempting to attack the argument and find fault with it, but 'critic' in a more positive sense. In this sense, a critic is someone who is attempting to analyze and evaluate the argument by some consistent and clear set of standards, and needs to do this by first of all identifying the argument by interpreting the text of discourse in which it has occurred. Such an interpretation typically involves clarification as well as explanation, but it is different from these other two speech acts in that it is an attempt to preserve the meaning of the original text of discourse, as opposed to changing it, in a way, for example, that might be required by clarifying it. Interpretation of an argument, in this sense, means offering an account of it that breaks the text of discourse down into a set of propositions in a key list, and trying to determine which inferences the author of the original discourse has postulated as relating some of these propositions to others. It needs to be noted that in the process of critical argumentation, often more than one interpretation needs to be considered. Such a multiplicity of interpretations can be very common in instances where the original text of discourse is vague, abstract, ambiguous, or unclear.

2   The three stages of critical assessment

The kind of work that makes logic useful as applied to real cases can be broken down into three parts or stages: identification, analysis and evaluation (van Eemeren and Grootendorst, 1984). The identification part is the task of determining what the argument is, supposedly, from the textual and contextual evidence given in the case. The basic questions to be asked here is: what are the premises and conclusions of the argument? If the argument has multiple conclusions, the job is to identify a longer sequence or chain of arguments aimed at some ultimate conclusion. The identification stage also concerns the identification of so-called missing premises and conclusions. At the identification stage, two things need to be accomplished. The first thing is to determine the text of discourse that is to make up the case. The second thing is the identification of a localized item within that text of discourse that is to be the focus of the evaluation. Normally it is an argument. But it could also be a question, an explanation, or some other kind of speech act that is connected to argumentation. Let's say that it is an argument that is the focus of evaluation. What needs to be done is to identify that argument within that text of discourse. Initially, the argument is identified by specifying its conclusion. When scanning a text of discourse, the conclusion is the first part of an argument to be identified. The conclusion is the claim that is supposedly being made. The analysis stage involves identifying the type of argument involved. For example, it could be an argument from analogy or an argument from appeal to expert opinion. These two stages are not independent, for finding missing premises may best be assisted by identifying the form of the argument required to make it structurally correct. The evaluation stage is one of determining whether the argument is strong or weak, correct or fallacious, by the structural standards appropriate for that type of argument.

The next thing is the identification of the explicit premises that have been put forward to support that conclusion. Thus an argument is a set of propositions. But it is more than that. It is a set of propositions as a sequence of reasoning. And it is reasoning used for some conversational purpose. Thus at the identification stage, one task is identification of the conclusion and premises that make up the reasoning in the argument. The other task is to judge what purpose that reasoning is supposedly being used for in the given case.

This latter task is contextual. One has to look at where the argument came from. Was it from a textbook, a newspaper editorial, a commercial advertisement, or a scientific source, for example? Who wrote it, and if it was an editorial from a newspaper for example, what was the topic of the editorial, as indicated by the headline, or other information. This contextual information is vitally important, for example, in judging whether the local argument in the given case is relevant.

As noted above, an important part of the analysis stage is the identification of implicit premises. This task is the analysis of so-called enthymemes, or argument with implicit premises or conclusions that need to be reconstructed from the text of discourse. Again, what is required as evidence is the contextual information about the supposed purpose of the text. Conclusions can be implicit, but typically what most of this work involves is the determination of implicit premises. The implicit premises are inserted as hypotheses on the basis of plausible assumptions about what the arguer supposedly or plausibly meant to say. As shown below, they are often conjectured by the respondent of an argument on the basis of common knowledge about how things can be expected to go in domains familiar to both arguer and respondent. As hypotheses, they need to be marked as separate from the explicit premises.

One of the main tools of the analysis stage is the argument diagram. The argument diagram identifies the whole chain of reasoning in a given case, from some initial premises to an ultimate conclusion. Each proposition is marked as a point in the diagram. The arrows joining the points represent steps of inference from one set of points to other sets of points. There can be multiple conclusions. But most often, the characteristic pattern one finds is that of multiple premises leading to a single conclusion. Then that conclusion may serve again as a premise in a next argument. The whole sequence in the given case is mapped out by the argument diagram.

Once the argument diagram has been constructed, and the missing premises filled in, the argument can be evaluated. The argument to be evaluated is generally some localized subargument that occurs as one single argument within the larger argument diagram. The first step of evaluation is to identify what type of argument it is supposed to be. Is it deductive or inductive, or does it fall into the third class of so-called abductive or plausible arguments? The third class, or even whether there is such a third class, is a matter of current controversy. Traditionalists think there are only two kinds of arguments, deductive and inductive. But many now say there is a third type of argument that is nether deductive nor inductive. It is a weaker type of argument. It is a guess that could be wrong, but is put forward as a supposition based on incomplete evidence. This kind of reasoning is based not on what is always the case, or what is mostly, or probably the case. It is based on what can generally expected to be the case in a familiar kind of situation. It is a kind of guess or supposition that is subject to defeat once more comes to be known about the particulars of a given case. Abductive reasoning is also known as inference to the best explanation.

The thesis put forward below is that the filling in of implicit premises is accomplished by abductive reasoning. Part of the process of identification of an argument is to fill in the explicit premises and conclusion of a given argument. Part of this abductive process is to look for a plausible interpretation of the argument.

If confronted by ambiguity, unclarity or vagueness, several possible interpretations may have to be considered. What the evaluator needs to do is to look for the most plausible interpretation. This judgment depends partly on the purpose and context of the discourse. What is often invoked at this stage is the so-called principle of charity. What the evaluator should do is to pick the interpretation that makes the argument in the given case seem strongest. Given a choice of two or more possible interpretations, the one that is the most plausible should be picked as representing the arguer's meaning. However, in some cases, the tie will be close. The best one can say is that either of two interpretations are plausible, or roughly equally plausible. What the evaluator needs to do in such a case is to work up two argument diagrams side by side. Then two evaluations can be given, depending on what the arguer supposedly meant. Below, it is shown that the principle of charity is not specific enough to give adequate guidance. What is necessary is for the critic to simulate the arguer's reasoning to get a plausible explanation of it.

Another common problem of interpretation is that of trying to decide in a given case whether a sequence of reasoning is an argument or an explanation. The solution to this problem is to realize that this judgment depends on the purpose of the discourse in the given case. The purpose of an argument is to try to prove (or disprove) some claim that is subject to doubt or controversy. The proposition at issue is in doubt. The purpose of an explanation is to throw light on some proposition that is presumed to be true. It is presumed to be a fact. It is not in doubt. The goal is to answer a question like "Why is it so?", or "How does it work?" These probing explanatory questions do not express doubt that the proposition in question is true. They ask for some account that can throw light to the questioner on how it came about. The difference is not in the reasoning itself. The difference is in how the reasoning was supposedly used for some purpose in a given case.

It is often said that the process of interpretation of a text of discourse is subjective, because a critic can never really get inside an arguer's head, and really know for sure what she meant to say. This remark is partly correct, because interpretation of a text of discourse is based on plausible reasoning. It is a guess, or hypothesis, about what the arguer supposedly meant, or can be taken to mean. But is it a hypotheses that can be based on verifiable or falsifiable evidence, external to the arguer's "head" or mind? The basic evidence is the text and context of discourse in the given case. From this objective evidence, hypotheses about how to take the argument can be constructed and evaluated. In many ways, the process, if properly carried out, is quite objective. It can be based on verifiable evidence, and on a process of logical reasoning that can be checked and evaluated by objective criteria. Of course, the kind of logical reasoning used in typical cases tends to be abductive. But abductive reasoning has objective standards and criteria for analysis and evaluation.

Of course, there can be many reasons for interpreting a text of discourse. The critical point of view adopted by logic is only one of many possible viewpoints that may be taken. Statutory interpretation has been acknowledged as a very important kind of evidence in law. Interpretation of discourse is very important in history, religion, and literature. Each of these different attempts to interpret discourse can have different purposes and different guidelines. Each has different ways of using textual evidence, and deciding what counts as good evidence for a claim. But the methods used to interpret discourse in law and religion are quite instructive and interesting to those of us trying to build an applied or informal logic for critical thinking.

3   Plan recognition and incomplete arguments

In the current logic textbooks, deductive logic, like the theory of syllogistic reasoning, is the usual method used to fill in missing premises and conclusions in incomplete arguments, or so-called enthymemes. This technique can be somewhat useful, but it distorts the given argument in many cases. In many common everyday arguments, the thread of reasoning is a means-end sequence of practical reasoning based on common knowledge shared by the arguer and the respondents on how things normally work in a situation familiar to both. Those of us who teach informal logic know that the syllogistic representation of incomplete arguments is stilted and artificial, and that a much more natural representation of the line of argumentation in such cases can be given. The problem is that the more natural representation seems to make the argument more complex, and there does not seem to be any formal system that could be used to generate the structure of the argument in a mechanical or automated way. But there is some source of help in dealing with this problem. Plan recognition technology has been developed to deal with similar problems of representing practical reasoning in natural language discourse. Plan recognition deals with examples of natural language discourse that are very similar to examples of incomplete arguments as studied in logic, and has developed some tools that can be applied to such examples of incomplete arguments.

In the TRACK system of plan recognition developed by Carberry (1990), a system of focusing is used. From a proponent's utterance in a dialog, the respondent infers the proponent's immediate goal. Then the proponent uses the domain-dependent knowledge base of goals and plans in the system to infer other goals and plans from the immediate goal inference. The immediate goal inference could be based on a Gricean implicature for example. Suppose, for example, an honors philosophy student asks a philosophy professor if she must take logic. The professor would infer that the student's immediate goal is to fulfill the requirements for a philosophy honors degree. But the professor and the student both have domain-

dependent knowledge of how the system of requirements works in a university setting. The professor would infer that the student's broader goal is to graduate from university. And they both share the knowledge that in order to graduate you have to take a certain number of courses, and that to graduate with an honors degree, there are rules that place restrictions on which courses need to be taken. This kind of knowledge is called domain-dependent, because it is knowledge about how a specific activity is performed or on some specific domain in which proponent and respondent are in a special position to know about. The TRACK system is called a focused system, because in plan recognition a proponent utilizes this shared domain-dependent knowledge to infer the goals and plans of a respondent.

The problem posed by typical cases of plan recognition is very similar to the traditional problem of so-called enthymemes in logic, or incomplete arguments. In plan recognition, the plan recognizer must take the given utterance, and try to make sense out it by filling in missing assumptions on what the speaker's goals and plans presumably are. This facet of the problem is called "ellipsis" in plan recognition. Filling in the missing parts of a speech act using ellipsis is typically a problem of finding common goals and plans based on common knowledge (so-called) concerning the ways that things are done in situations familiar to the speaker and hearer. A good example of an ellipsis problem is provided by the following short dialog from (Flowers and Dyer, 1984).

> *Proponent*: The Korean jet shot down by the Soviets was a spy plane.
>
> *Respondent*: With 269 people on board?

The question posed by the respondent is more than just a request for information. As Carberry (1990) pointed out, to really interpret what the question is saying in an adequate way, an analyst has to "recognize the surprise and doubt conveyed by the elliptical fragment". The respondent is expressing doubt about the assertion that the Korean jet was a spy plane. The basis for recognizing the thrust of the question is the unstated assumption that sending a would-be spy plane over a foreign country with 269 people on board would be absurd. Why? The reason is based on common knowledge about how the activity of spying using aircraft normally works. We know that it is done by an aircraft containing cameras and other automated technical devices, requiring a pilot and perhaps a small crew needed to navigate the plane and monitor the technical devices. We know that it is generally a risky activity, because the plane flies over the airspace of a foreign country, and it is possible that it can be intercepted or even fired on. Incidents have been reported in the past where spy planes have been shot down. Hence using a plane with 269 people on board would not only be unnecessary. It would also pose an unnecessary risk of loss of life for no apparent reason.

The respondent's speech act in the dialog above syntactically has the form of a question. But probing deeper beneath the surface, the question is, pragmatically speaking, being used as an argument. It is used to express doubt. It is used to argue that the assertion that the Korean jet was a spy plane is implausible. But what is the evidence used to back up this argument? It is quite persuasive evidence, based on implicit assumptions about how things are normally done in a domain of knowledge on how an activity is normally done. This domain of knowledge is specific, but one would expect it to be shared by the proponent and the respondent, and those of us who attempt to analyze the utterance using techniques like plan recognition. This example shows how similar are the problems and methods in both plan recognition and analysis of enthymemes or incomplete arguments in logic. In trying to determine the unstated premises in an argument, the analyst must typically extract statements that are presumed to be part of the argument, on which the argument rests as assumptions, but that have not been explicitly stated by the proponent. The task is closely similar to that posed by the problem of ellipsis in plan recognition. Both tasks are based on underlying unstated assumptions that the given argument (or other speech act) depends on for its force and effectiveness. In both instances, these assumptions are based on common knowledge about the ways things are normally done, typically in a specific domain known to both the speaker and the hearer in a dialog.

In his widely used logic textbook, Hurley (2003, p. 279) put the following argument in a set of exercises for the student. The student is instructed to supply the missing premise or conclusion, and to attempt whenever possible to convert the enthymeme into a valid argument.

The financial aid argument

College students of today are the higher income taxpayers of tomorrow. Congress should consider financial aid as an investment in the financial future of our country.

The first statement is a premise, and the second is the conclusion of an argument. What are the missing parts that should be inserted to make the argument complete, or at any rate, less incomplete? The answer given by Hurley in the back of the textbook (p. 607) is to reformulate the argument as the following two syllogisms.

Syllogism 1

All policies that promote more college graduates tomorrow are policies that result in higher tax revenues tomorrow.

All policies that offer financial aid to college students today are policies that promote more college graduates tomorrow.

All policies that offer financial aid to college students today are policies that result in higher tax revenues tomorrow.

Syllogism 2

All policies that result in higher tax revenues tomorrow are good investments in the future.

All policies that offer financial aid to college students today are policies that result in higher tax revenues tomorrow.

All policies that offer financial aid to college students today are good investments in the future.

This analysis is somewhat helpful to beginning logic students, as it helps them to see that unstated premises are part of what makes this argument convincing. Thus it shows students how hidden assumptions in an argument can be identified and questioned. A drawback is that if you interpret the universal quantifier 'all' as meaning 'all without exceptions', each premise is arguably false. The least plausible premise is 'All policies that result in higher tax revenues tomorrow are good investments in the future.' This statement is politically controversial, to say the least. By representing the argument as having a premise that is so highly questionable, or even plausibly false, there is danger that this representation of the argument could be accused of committing the straw man fallacy. The question is thus raised: is there some way of representing the reasoning of the argument as otherwise than deductively valid, without positing what strongly seem to be false premises? There is another analysis that can show more naturally how various unstated assumptions are linked to the given premise in order to lead to the conclusion by a chain of reasoning.

There are various missing assumptions that should be cited to begin with. One is that 'financial aid' means financial assistance to college students. Another is that higher income taxpayers as a group contribute significant funding to the government. Of course, some higher income taxpayers escape taxation by various means. But generally speaking, subject to exceptions, higher income taxpayers pay more money in taxes. Another is that if such a group is enlarged in the future, producing more taxes for the government, it would make the financial future of the country better. Another is that making the financial future of the country better is a good thing, a worthy goal. The whole argument hangs together as a sequence of means-end practical reasoning. Starting with the given premise, various unstated premises can be added to make this sequence of reasoning more explicit, showing how the sequence leads to the conclusion. The first two statements, for example, are not seen as universal generalizations made false by one exception. They should be seen as rough generalizations that are subject to exceptions, and are not falsified by a

single exception. They make statements about the way things can normally be expected to go in a situation familiar to speaker and hearer.

A better analysis of the financial aid argument can proceed by, first of all, breaking the argument down into a sequence of propositions.

Sequence of propositions in the financial aid argument

> Explicit Premise: College students of today are the higher income taxpayers in the future.
>
> Implicit Premise 1: Under our system of taxation, those who earn higher incomes tend to pay higher taxes.
>
> Implicit Premise 2: If a larger number of citizens pay higher taxes, government revenues are increased.
>
> Implicit Premise 3: If government revenues are increased, the financial future of the country will be made better.
>
> Implicit Conclusion 1: Increasing the number of higher income taxpayers there will be in the future will make the financial future of the country better.
>
> Implicit Premise 4: giving financial aid to college students now will increase the number of college students who go on to graduate and then, in the future, earn higher incomes.
>
> Implicit Conclusion 2: Giving financial aid to college students now will increase the number of high income taxpayers there will be in the future.
>
> Implicit Premise 5: Anything that contributes to the goal of making the financial future of the country better is an investment in the financial future of our country.
>
> Implicit Premise 6: Giving financial aid to college students will contribute to the goal of making the financial future of the country better.
>
> Explicit Conclusion: Congress should consider financial aid as an investment in the financial future of our country.

This argument depends for its convincingness on several unstated but implied links between the means and the end. In particular, it implies that there is some sort of practical linkage between making the financial future of the country better and giving financial aid to students now. The implied linkage is one that relates to the way things normally work in a government in matters of taxes and finance. Our government, or any democratic system government of the kind we are familiar with, gets most of its funding from taxes. But not everybody pays the same level of taxes. The tax system is such that higher income citizens pay higher taxes. That statement is part of common knowledge, and both the proponent and the respondent know that

our tax system works in this way, by graduated levels of income. Thus generally, if you have a higher income, you will pay higher taxes. Of course, this generalization, like the others above, is subject to exceptions. Nonetheless, they can be used as premises in practical reasoning as long as we realize they are subject to default.

The practical reasoning method of representing incomplete arguments is not only much more natural in a broad range of cases. It can be seen where the missing statements come from. Both the proponent of the argument and the readers (students, respondents) share common knowledge of how our system of taxation works, and how it generally means that those in the higher tax brackets will tend to pay larger amounts into government revenues. How our system of taxation works in these commonly known respects can be seen as domain-dependent knowledge. In this case, it is not specialized technical knowledge but what is often called common knowledge. It is really a kind of familiarity with how things generally work in an everyday domain known to the participants in an argument. To fill in the missing premises, an argument analyst needs more than just deductive logic. What is needed is practical reasoning and special domain knowledge of how things generally work on some subject, like our system of taxation.

What has been shown is relevant to the dispute that surfaces from time to time in informal logic on whether you need knowledge of the subject-matter to critically analyze arguments. This contention has been posed as a threat to the whole enterprise of informal logic. For if teaching critical thinking presupposes knowledge of domain-dependent subject matter of arguments, it would need to be taught in a different way. Thus teaching critical thinking in chemistry, for example, might use quite different methods from teaching critical thinking in history. But what has been shown above is that even though analyzing incomplete arguments does depend on domain-dependent knowledge, it can still be carried out by a generalizable technique while allowing for this dependency as applied to specific cases. The practical reasoning structure can be applied to a broad range of cases in the same way, except that focusing on the more fine-grained linkages in the sequence of argumentation involves domain-dependent knowledge shared by speaker and hearer. This domain-dependent knowledge can however, in many instances, simply be common knowledge about the ways familiar sequences of events or actions can normally be expected to go, subject to exceptions.

4   New tools for argument diagramming

The task of interpreting a given text of natural language discourse prior to evaluation of the arguments expressed in it cannot be carried out in a purely mechanical way. The reason is that the discourse interpreter needs to share a grasp of how the

author of the discourse is using that discourse. Is she using it to make a claim, and support that claim by giving reasons to accept it? Or is she simply trying to get the reader (audience) to come to understand something that might have been unclear to them? To decide such an issue in a concrete case, evidence, like the presence of so-called indicator words, is helpful (Snoeck Henkemans, 1992). But in the end, the shared understanding of speaker and audience about how things normally go is the ultimate basis of how the text should be interpreted. Of course, this shared understanding is implicit in, and dependent on the aims of the conversation the speaker and hearer are presumably engaged in, as Grice showed. Despite these reservations about the task of identifying and analyzing arguments in natural language discourse, some tools are proving to be very useful to assist in the task. One of them, as indicated several times previously, is the method of argument diagramming. The diagram gives at least a provisional model of the structure of an argument that can be used to assess various problems of evaluation. For example, an argument diagram can be extremely useful in helping to amass evidence on the question of whether a lengthy and complex chain of argumentation is circular.

One automated system of argument diagramming is the *Araucaria* software developed by Chris Reed and Glenn Rowe of the University of Dundee. The *Reason!Able* system is very simple to use and learn, and thus seems naturally applicable to the teaching of critical thinking. The *Araucaria* system has been designed more with uses in the theory of argumentation in mind. It is a more powerful system that gives direct capability for carrying out several tasks of argument reconstruction that have long been stressed as useful in logic textbooks and manuals. *Araucaria* represents linked and convergent arguments differently, and it helps the user find missing premises and conclusions in incomplete arguments, or so-called enthymemes. *Araucaria* allows the user to mark up arguments, by assigning numbers to the premises and conclusion, once these statements have been identified in a given text by the user. *Araucaria* can also help the user find argumentation schemes in common arguments. Based on such a prior partial identification, *Araucaria* can then identify the missing premise or conclusion. For example, suppose the user finds that sentence 1 is the conclusion of an argument from verbal classification, and that sentence 2 is the instantiation of the first premise of that argumentation scheme. Then *Araucaria* can fill in the missing premise by using the argumentation scheme to match the instantiation. The ability to identify argumentation schemes is also very important for other purposes in the analysis and evaluation of argumentation. Argumentation schemes are often very important both in helping to distinguish between linked and convergent arguments, and generally in helping a user to grasp the structure of an argument and pose appropriate critical questions.

Araucaria can load either a file containing a text of discourse or an argument that has already been "marked up" by having numbers assigned to each premise and conclusion that has been identified by the user. This information appears on the left side of the screen. Selecting any part of this text makes the right number of nodes (circled numbers) appear on the right side of the screen. Then the user can drag a line from one node to another, indicating a support relationship. Multiple lines of support can be joined together by a distinctive type of bracketing line indicating a linked argument. Otherwise, drawing a line from one node to another indicates a convergent argument. In short, *Araucaria* produces a typical looking argument diagram in the Freeman style, with interconnected linked and convergent arguments. As noted above, it can also recognize common argumentation schemes of the kind in (Walton, 1996), and thus use these to find missing (unstated) premises and conclusions in an argument. Once the analysis has been completed, resulting in an argument diagram, both the original text of discourse and the diagram can be saved. Thus *Araucaria* is very simple to use, and fits the usual format of the technique of argument diagramming recommended in most logic textbooks.

The technical features of *Araucaria* have been described in (Reed and Rowe, 2001). The system is based on the tree structure in which one ultimate conclusion is identified (the root point of the tree), and all the lines in the diagram lead into that node. The argument markup language defines tags that indicate components of arguments (statements), support relationships among the statements, and instances of argumentation schemes. The markup language is based on XML (Extensible Markup Language), a widely used industry standard.

The decisive advantage of *Araucaria* over other software method of argument diagramming is that it can use its storehouse of common argumentation schemes to help the user fill in missing premises in incomplete arguments. For example, practical reasoning has been identified in (Walton, 1996, pp. 11–12) as an argumentation scheme, with a matching set of critical questions. *Araucaria* therefore is a system that has the capability to fill in the missing assumptions in a typical case like the argument from Hurley above about giving financial assistance to students. A system like *Araucaria*, that can have common argumentation schemes built in, could help the user to pick out the missing premises in this argument by using the argumentation scheme for practical reasoning to identify the missing parts needed to make the argument fit this pattern. Not only that, *Araucaria* can recognize that practical reasoning represents a defeasible form of reasoning that is different in its structure and requirement from typical deductive forms of reasoning like syllogistic arguments. Through its capability to represent defeasible reasoning, based on argumentation schemes, *Araucaria* has a striking advantage over other systems of automated argument diagramming.

To give the reader an idea of how *Araucaria* works, an analysis of the financial aid argument presented above is represented below in an *Araucaria* argument diagram. Normally in such a diagram, implicit premises or conclusions are represented as statements that appear in a darkened box with a broken border. However, to make the argument diagram for this example as simple and easy to appreciate as possible both implicit and explicit premises and conclusions are marked as such in the boxes on the diagram. If the reader looks back to the set of statements identified in the previous section that represent all the premises and conclusions in the financial aid argument, it can be seen by looking at figure 7.1 how each of these statements is related to the others in the argument as a whole.
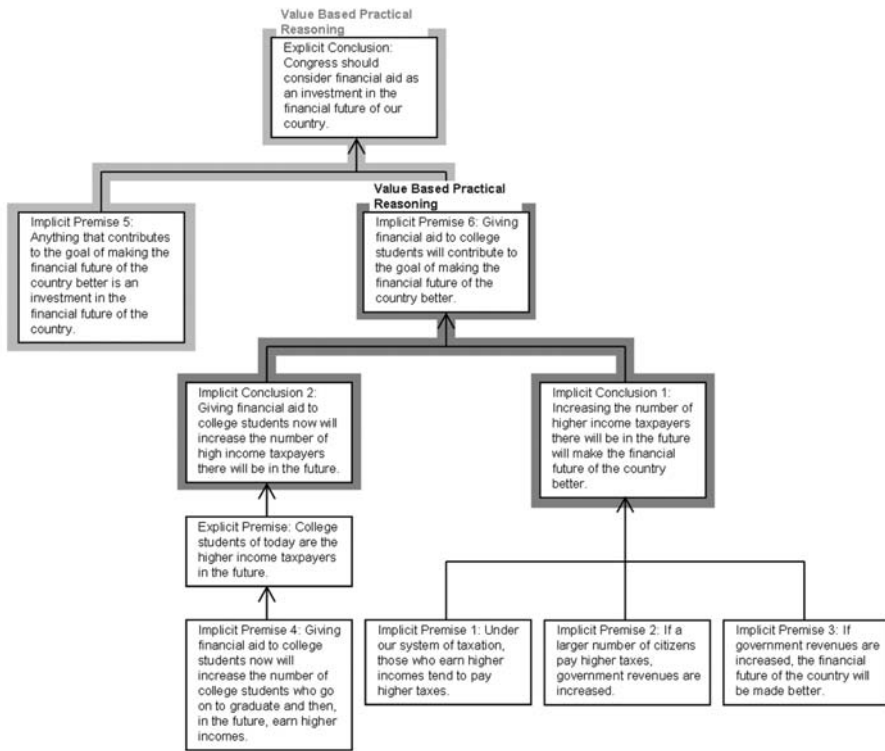
Figure 7.1 *Araucaria* argument diagram of the financial aid argument

Analyzing the financial aid argument as shown in figure 7.1 shows how the scheme for value-based practical reasoning described in chapter 4 (section 1) can be used to link premises together in a chain of argumentation. Note that it is arbitrary, in some instances, whether an implicit statement in an argument is designated as an

implicit premise or an implicit conclusion. The reason is that when the component statements are chained together to make a longer argument, the statement that is a conclusion in one argument can also function as a premise in the next argument.

5   The problem of enthymemes

An enthymeme, according to the conventional meaning of the term in logic textbooks, is an argument that contains a missing (not explicitly stated) premise or conclusion that needs to be considered as part of the argument in order for it to be a defensible and convincing argument of the kind it was evidently meant to be. The classic example is the argument, 'All men are mortal, therefore Socrates is mortal'. It is assumed that Socrates is a man, and the argument is valid once this missing premise is added. And hence that it is reasonable to accept that premise as part of the argument. The financial aid argument in the previous section is a textbook example of an enthymeme, in this instance one that has many missing premises. It is a widely acknowledged problem with enthymemes (Burke, 1985; Gough and Tindale, 1985; Hitchcock, 1985) that if a critic is given carte blanche to fill in any proposition needed to make an argument valid, she may be inserting assumptions into the text of discourse that the speaker doesn't accept, or would even disclaim. The question is how we can be sure that a statement selected as a missing premise or conclusion in an argument really is one.

The first step toward sorting out the problem is to draw the distinction, following Ennis (1982, pp. 63–66) between needed and used assumptions. *Needed assumptions* in an argument are missing propositions such that (1) the argument is not structurally correct as it stands, but (2) when the propositions in question are inserted, the argument becomes structurally correct. *Used assumptions* in an argument are propositions that, even though not explicitly stated in the text of discourse, are meant to be part of the argument by the speaker, and are likely to be so taken by the hearer. The difference, according to Ennis (1982, p. 64) is that used assumptions can be taken as unstated reasons supporting an argument, while needed assumptions may or may not be.

Filling in used assumptions is a hard task in some cases, because selecting missing premises depends on determining what an arguer meant to say as part of her argument. In many case, an arguer may be unclear on what she means, and may even claim she is not committed to some assumption inserted as a missing premise argument. The omnipresent danger is that of committing the *straw man fallacy*, the fallacy of attributing a proposition to a speaker's argument that exaggerates or distorts the argument in order to make it easier to refute (Scriven, 1976, pp. 85–86). Johnson and Blair (1983 p. 71) defined the straw man fallacy as being

committed "when you misrepresent your opponent's position, attribute to that person a point of view with a set-up implausibility that you can easily demolish, and then proceed to argue against the set-up version as though it were your opponent's." The straw man fallacy can involve exaggeration as well as other forms of distortion. The following example of a dialog (Freeman, 1998, p. 88) can be used to illustrate this fallacy.

The beer and wine example

> CONCERNED CITIZEN: It would be a good idea to ban advertising beer and wine on radio and television. These ads encourage teenagers to drink, often with disastrous consequences.
> ALCOHOL INDUSTRY REPRESENTATIVE: You cannot get people to give up drinking; they've been doing it for thousands of years.

Freeman (p. 88) observed that there is no evidence that the concerned citizen is maintaining that teenagers or other people should give up drinking. But there is evidence that the alcohol industry representative wants to make us believe that the concerned citizen advocates this position.

To analyze the straw man fallacy in this case, Freeman contrasted the following pair of propositions (p. 88) posing the question: which is the easier one to refute?

> A: It would be a good idea to ban advertising beer and wine on radio and television (the concerned citizen's original conclusion).
> B: It would be a good idea to get people to stop drinking (the alcohol industry representative's portrayal of that conclusion).

B is much easier to refute than A. On this assumption Freeman (1988, p. 88) argued that the alcohol industry representative has misrepresented the concerned citizen's position, "in a way making it easy to refute, making it look almost silly". On this basis, he classifies the argument as an instance of the straw man fallacy.

One way of coping with the problem is to invoke the *principle of charity*, which requires that, given a choice of selected assumptions that could be added as a premise or conclusion, pick the one(s) that makes the argument stronger. Trying to use this principle, however, introduces two additional problems (Gough and Tindale, 1985, p. 102). The first is the problem of trying to decide whether or not the evaluator has produced a new argument to support the conclusion, as opposed to the original argument. The second is to determine how many premises should be required to produce the best possible argument out of the given text. For depending on the depth of analysis the critic wants to achieve, more premises can often be added.

A method widely used in logic textbooks is to apply deductive logic to the given argument in order to see what missing propositions are needed as additional premises (or as a conclusion) that would make the argument deductively

valid. For example, Hurley (2000, p. 289) defined an enthymeme as "an argument that is expressible as a categorical syllogism but that is missing a premise or a conclusion." He used the following example: "The corporate tax should be abolished; it encourages waste and high prices." (p. 289). This argument has the unstated premise, 'Whatever encourages waste and high prices should be abolished'. Once that missing assumption is inserted as an additional premise, the argument becomes deductively valid.

It can be argued that not all enthymemes are based on deductive logic, and that many of them are better analyzed as being based on defeasible generalizations and defeasible argumentation schemes (Walton and Reed, 2005). Consider the generalization that whatever encourages waste and high prices should be abolished. There are a lot of things, like manufactured products of many kinds, that encourage waste and high prices, but nevertheless it might be argued that even so, they ought not to be abolished.
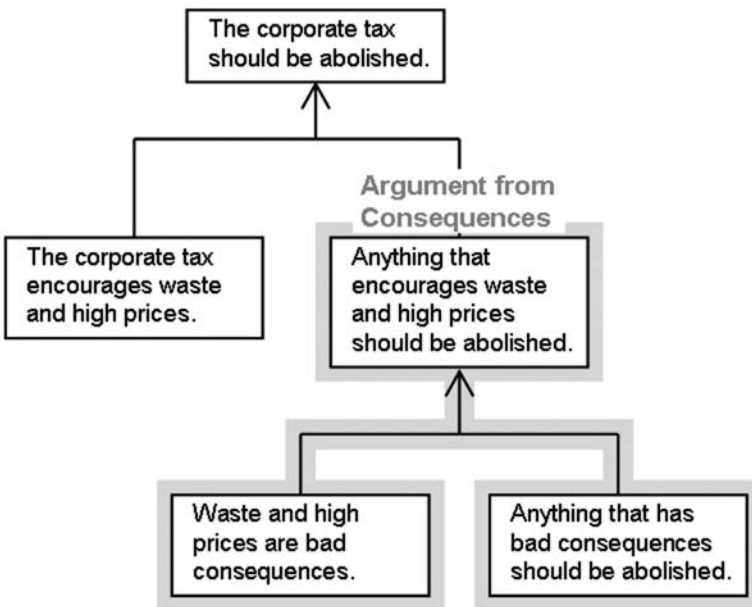
Figure 7.2 Argument diagram of the corporate tax example

Here we see the use of an argumentation scheme representing a defeasible type of argument, argument from negative consequences. But where does the premise 'Waste and bad consequences are undesirable' come from? It can be argued that it

can be classified under the heading of common knowledge. In the next section, it is argued that there are three bases for the enthymeme (1) filling in implicit premises or conclusions that are needed to make the given argument satisfy some standard of structural correctness like deductive validity, (2) filling in implicit premises or conclusions based on common knowledge, and (3) filling in implicit premises or conclusions that can reasonably be taken to represent commitments of the arguer who put the given argument forward.

6   Three bases for the enthymeme

According to Govier (1992, p. 120), an assumption in an argument is based on common knowledge if it states something that is known by virtually everyone, even though such matters are dependent on audience, context, time and place. She cites the examples "Human beings have hearts" and "Many millions of civilians have been killed in twentieth-century wars" (p. 120). Freeman (1995, p. 269) cites common knowledge as a basis for premise acceptance in arguments by classifying it as a form of presumption based on the shared "lived experience" of a speaker and hearer (p. 272). It seems reasonable that common knowledge, as defined by Govier and Freeman, could be a rationale for filling in missing premises or conclusions in enthymemes. Common knowledge has also been recognized as a way of filling in gaps in reasoning in artificial intelligence studies. Everyday reasoning of a kind often necessary to represent in computing is based on common implicit knowledge that all human beings share. For example, we all know that if President Bush is in Washington then his head is also in Washington. The open mind common sense system (OMCS)[1] is a common sense knowledge acquisition system that collects items of common knowledge like the following propositions (Singh, Lin, Mueller, Lim Perkins and Zhu, 2002, p. 3).

> People do not like being repeatedly interrupted.
>
> If you hold a knife by its blade then it may cut you.
>
> People pay taxi drivers to drive them places.

These three statements are defeasible generalizations, as opposed to absolute universal generalizations of the kind that are used in deductive logic.

As shown by the financial aid argument analyzed above, in many cases of everyday argumentation, enthymemes depend on premises that are implicit defeasible generalizations that are tentatively acceptable because they represent common

---

1.   http://commonsense.media.mit.edu/cgi-bin/search.cgi

knowledge about the way things commonly work in everyday life. To cite another example (Walton, 2001), we all know as common knowledge that soup is generally eaten with a spoon, and not a fork. Consider the following example of an argument quoted from (Acock, 1985, p. 106).

The risi e bisi example

> Risi e bisi is often listed on menus among the soups, and some gastronomic writers dare to call it one. Nonsense! It is served with a fork. Who ever heard of eating soup with a fork? ("Risi e bisi," *The Best of Italian Cooking*, by Waverly Root, p. 219).

An explicit premise of the argument is that risi e bisi is served with a fork. The rhetorical question, "Who ever heard of eating soup with a fork?" can be assumed to make the statement that nobody eats soup with a fork (as a general practice). If we analyze the argument this way there is an implicit premise based on common knowledge: if something is served with a fork, and nobody eats soup with a fork, then what was served is not soup. Once this premise has been made explicit, an implicit conclusion can be drawn: risi e bisi is not a soup. To find the unstated assumptions, we need to have the common knowledge that soup is generally not eaten with a fork, and that a restaurant will generally try to furnish a diner with the appropriate utensil for eating. But we also may assume, based on common knowledge, that eating soup with a fork would not be practical, and thus we can derive the conclusion that if something is served with a fork, it is not being treated as soup.

These examples and others show that common knowledge is an important basis for making the implicit premises or conclusions in an enthymeme explicit. However, it can be argued that common knowledge, in this sense, is not knowledge, in the way that term is used in epistemology. Instead it is more like what Aristotle called an *endoxon*, a commonly accepted opinion that can be taken for granted as part of an argument because no parties to the dialog would be likely to contest it. To handle examples of enthymemes, the dialog system CBV defined in chapter 5, section 4, needs to be extended to accommodate common knowledge. In any dialog, there is an opening stage, an argumentation stage, and a closing stage. At the opening stage, where the participants have agreed to follow an appropriate set of rules for rational argumentation, there needs to be agreement about what can be taken for granted as common knowledge. For example, all the participants might agree that any statement in an encyclopedia counts as common knowledge. Such statements are sometimes called common starting points (van Eemeren and Grootendorst, 1992). As the argumentation stage proceeds, once a statement has been identified as common knowledge, it can henceforth be treated as a commitment of all parties to the dialog. What is needed is an extension of CBV called CBVCK. In CBVCK, any statement agreed to be common knowledge

by both participants by the opening stage counts as an implicit commitment of all parties. As such, it can function as an implicit premise or conclusion in any chain of argumentation put forward by any party to the dialog.

The next example is an argument with two missing premises.

> Animals in captivity are freer than in nature because there are no natural predators to kill them.

The conclusion is the statement that animals in captivity are freer than in nature. The explicit premise is the statement that there are no natural predators to kill animals that are in captivity. One missing premise is the statement that there are natural predators to kill animals that are in nature. Another missing premise is a conditional: if animals are in a place where there are no natural predators to kill them, they are freer than if they are in a place where there are natural predators to kill them. This second implicit premise is also a generalization. The first missing premise is based on common knowledge, but the second one is highly controversial. It is a proposition that runs counter to common knowledge, or to common sense statements about what is generally accepted. Yet the arguer does seem to be committed to this argument. Thus in the argument diagram in figure 7.3, the first missing premise is labeled as common knowledge (CK), while the second is labeled as commitment (COM), meaning a kind of commitment other than common knowledge.
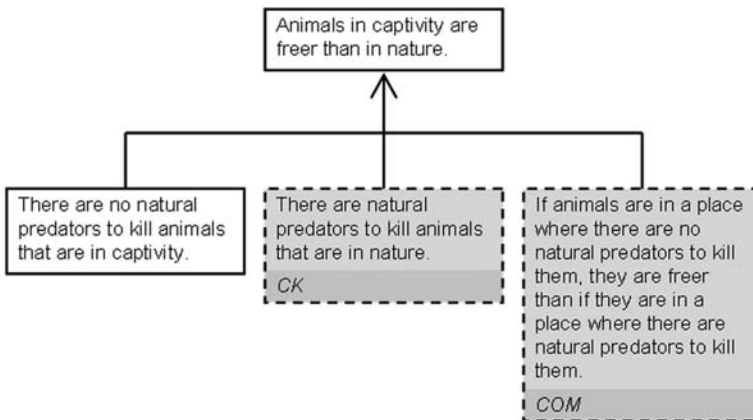
Figure 7.3 Types of implicit premises in the free animals example

We are given no further information about the text of discourse in which this argument appeared, but it does seem that the arguer is advocating a position against

animal activists who argue that keeping animals in captivity is a bad policy. He does seem to be committed to the assumption that if animals are in a place where there are no natural predators to kill them, they are freer than if they are in a place where there are natural predators to kill them. This assumption is based on a construal of the term 'free' that represents the arguer's own special position, one that differs from the generally accepted opinion. It needs to be classified as a missing premise based on the arguer's special commitment, as opposed to common knowledge that is generally accepted.

Thus the third basis for the enthymeme is commitment. Missing assumptions in an argument cannot always be elicited mechanically as statements needed to make an argument valid, nor are they always implicit assumptions based on common knowledge. In some cases, they are missing premises or conclusions that an arguer is committed to, as shown by other statements she has made or arguments she has put forward in a dialog. But to know which propositions are really commitments of an arguer, we come back to the notion of commitment in dialog, and the problem of how an arguer's commitments can be fairly determined by another party using the text of discourse representing the arguer's views. What is needed is a search tool that goes through the given text representing the arguer's previous moves in the dialog, and uses this text as data to judge whether a given proposition can reasonably be judged to be a commitment expressed in it or not.

Using of the commitment search tool outlined below assumes there are two parties (agents) engaged in a dialog of a specific type, like a critical discussion, A1 and A2. Each agent has a commitment set, and a search engine that it can use both to search through its own commitment set and that of the other party. A simple prototype search engine E1, that searches through a commitment store to see if a specific proposition is in it, is described. The target proposition the engine is searching for is designated 'P*'. If the engine searches through the whole set, and finds no proposition there that matches P*, it answers, 'The agent is not committed to P*'. Let us call the participant (critic) who is trying to determine what an arguer's commitments are A1, and the arguer who has put forward an argument in a dialog A2. Suppose A2 is committed to the proposition P1 and also to the conditional proposition, 'If P1 then P*'. A2 has never explicitly said that P* is true, but we could say that he is implicitly committed to P*. This could be shown by applying *modus ponens*. Thus A2 is implicitly committed to P*. The first type of search engine, prototype engine E1, goes through all the propositions in A2's commitment store and applies the rules of inference accepted by both parties in the dialog to subsets of them, singletons, pairs, and so forth. If it turns up a proposition P* by this method, P* can be said to be an implicit commitment of A2. If it fails to turn up proposition P* by this method, P* is said not to be an implicit commitment of

A2. Attributing P* to A2, and using this attribution to try to refute an argument of A2's, is committing the straw man fallacy.

Suppose you ask the search engine whether A1 is committed to P*. It starts with the first proposition, P1, it finds in A1's commitment set. X and Y are variables for propositions P1, P2,..., Pn that are commitments of an agent.

> Step 1: Apply each rule to each proposition in the commitment set, one at a time, and test to see whether P* follows from P1.
>
> Step 2: If P* is not yet determined as a commitment, move to the next proposition in the commitment set, P2, and apply steps 1 and 2 to it.
>
> Step 3: If P* is not yet determined as a commitment, move to the first pair of propositions in the commitment set, P1 and P2, and apply steps 1 and 2 to them.
>
> Step 4: Each time a new implicit commitment is found, add it to the original set.
>
> Step 5: Carry out this procedure recursively, until all the propositions and pairs of propositions in A1's commitment set are exhausted.
>
> Step 6: If P* is found, answer, "A1 is committed to P* because P* follows from commitment X, or from commitments X and Y".

When the search engine follows this list of instructions, it is looking for any indirect commitment that follows from any other proposition, or pair of them, that A1 is explicitly committed to. It does this by applying each of the rules of inference in the dialog successively to all singletons and pairs of propositions in A1's commitment set.

How could the search engine be applied to the beer and wine example? In this example, there is only a small text of discourse representing the two sentences in the dialog expressed by the concerned citizen. The concerned citizen claims that it would be a good idea to ban advertising beer and wine on radio and television because these ads encourage teenagers to drink. The third proposition that the concerned citizen asserts is that the ads encouraging teenagers to drink often have disastrous consequences. The alcohol industry representative commits a straw man fallacy when he attempts to refute this argument by improperly attributing the statement that it would be a good idea to get people to stop drinking to the concerned citizen. He improperly attributes this commitment to the concerned citizen because there is no evidence in the dialog that the latter has committed himself to this proposition. In this case, the task for the search engine is an easy one. It scans through the three propositions asserted by the concerned citizen, and applies all the rules of inference or argumentation schemes, that both parties in the dialog accept, to these three propositions. Through the kind of search process outlined above, it is not possible to come up with the proposition 'It would be a good

idea for people to stop drinking' as one of the commitments of the concerned citizen expressed in his argument. Therefore, when the alcohol industry representative uses this proposition to refute the concerned citizen's argument by supporting it with argumentation, he is committing the straw man fallacy. How the fallacy has been committed in this case can be shown by the argument diagram in figure 7.4.
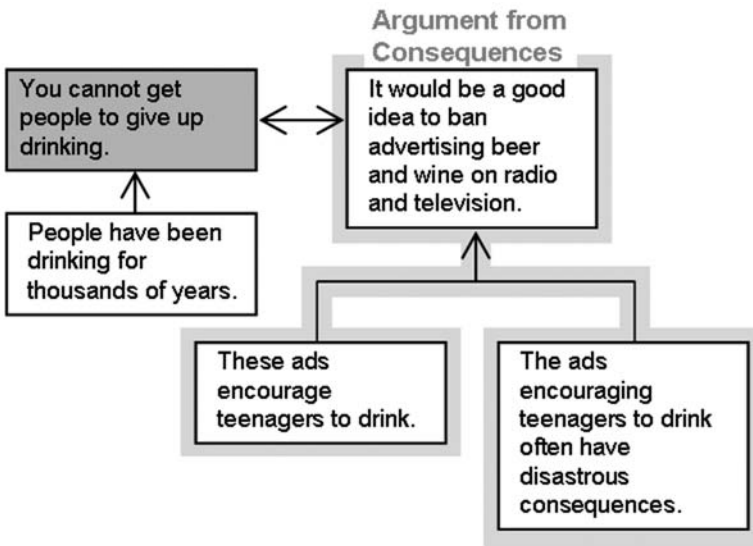
Figure 7.4 Argument diagram for the beer and wine example

The commitment search tool shows that the concerned citizen is not committed to the proposition that total abstinence is a good idea. By trying to refute the concerned citizen's claim that it would be a good idea to ban advertising beer and wine on radio and television by arguing that you cannot get people to give up drinking, the industry representative is implying, however, that the concerned citizen is committed to a policy of total abstinence. What textual evidence is there that the alcohol industry representative is trying to get us to believe that the concerned citizen is advocating total abstinence? The evidence is that the alcohol industry representative's statement, "You cannot get people to give up drinking", when placed in the dialog as his response to the prior move on the concerned citizen, shows that it is clearly meant to be a refutation move in the dialog. This refutation move is represented by the double arrow in figure 7.4 joining the industry representative's counterclaim to the thesis advocated by the concerned citizen.

The beer and wine example is an easy case, because the text of discourse is very small, and it is quite clear that the industry representative's argument is a straw

man fallacy. As we all know, however, there can be lots of hard cases where it can be a very difficult job to fairly judge whether some proposition is an implicit premise or conclusion in the argument contained in the given text.

7   Textual interpretation as an abductive process

In abductive reasoning, or inference to the best explanation, a hypothesis is suggested that helps to explain given data, like the material presented as the data in a text of discourse. The interpretation is based on a kind of hypothesis drawn by an inference to the best explanation of the text. The interpreter may construct several hypotheses, each of which can serve as an explanation of what is there in the given text of discourse, and then choose the best, or most plausible of these explanations. The reasoning is abductive because it is a kind of reasoning backwards from the given material in the text, to a conclusion that a particular account serves to explain the meaning of the text, or to bring out what is implicitly there. Once we come to be aware of the special abductive nature of techniques like argument diagramming, we can put the tasks of argument identification and analysis into a better perspective. We can see that an argument diagram represents a hypothesis that can be based on evidence, and that alternative interpretations can also be possible, also based on evidence. This perspective begins to make sense, showing that argument interpretation is not "subjective", in a sense meaning that it has no logical structure on which it is based, once abductive reasoning is seen to have a structure.

A very clear account of the structure of abductive reasoning has been given by Josephson and Josephson (1994, p. 14). By this account, an abductive inference has the following form, where H is a variable representing a hypothesis and D is a variable representing a given set of data or (presumed) facts.

Form of Abductive inference (Josephson and Josephson)

> D is a collection of data.
>
> H explains D.
>
> No other hypothesis can explain D as well as H does.
>
> Therefore H is plausibly true.

But how should an abductive inference be evaluated? According to Josephson and Josephson, the judgment of plausibility associated with an abductive inference should be evaluated differently from comparable evaluations of deductive or in-

ductive reasoning. Evaluating inductive reasoning is partly contextual, and depends on the following factors (p. 14).

1. how decisively H surpasses the alternatives
2. how good H is by itself, independently of considering the alternatives (we should be cautious about accepting a hypothesis, even if it is clearly the best one we have, if it is not sufficiently plausible in itself)
3. judgments of the reliability of the data
4. how much confidence there is that all plausible explanations have been considered (how thorough was the search for alternative explanations)

Two additional considerations required for the evaluation of an abductive inference (p. 14) are explicitly contextual in nature.

1. pragmatic considerations, including the costs of being wrong, and the benefits of being right
2. how strong the need is to come to a conclusion at all, especially considering the possibility of seeking further evidence before deciding.

The conclusion to be inferred in an abductive inference is the "best" explanation of the given data. However, such a judgment of what is best is relative to the given data. An abductive inference results in a conclusion that is typically a plausible assumption. It may have to be withdrawn in the future, because new data may come into a case. Collection of new data may indicate an explanation that was less plausible than the one previously accepted is now more plausible.

A good example of abductive reasoning is scientific reasoning at the discovery stage, or early hypothesis construction stage of scientific research. It is a process of guesswork that proceeds by constructing hypotheses that would explain the given data. Then a best or most plausible hypothesis is selected out that is worth testing. Abduction saves money and effort by picking out the most plausible hypothesis instead of experimentally testing all of the available hypotheses. There is a parallel between abductive reasoning in science and abductive reasoning of the kind used in textual interpretation. Interpretation of a text of discourse proceeds by constructing an argument diagram, or some comparable representation of what the critic takes to be the argument. Just as in scientific reasoning, the evidence is provided by the given data. But in the case of interpretation of a text of discourse, the evidence is the actual text of discourse in the given case. So although an interpretation of a text of discourse is conjecture, it can be seen as a hypothesis that is based on evidence that can be cited, and can be verified by other parties.

Taking the viewpoint of critical argumentation, we have in this chapter analyzed several examples of arguments, and in particular we have chosen examples of enthymemes to illustrate how such a process of analysis requires interpretation by finding implicit premises and conclusions in a given argument. Typically this

process can be described as working along the following lines. The critic scans over the text and finds some part of the text that looks like it could be an argument. However, in order to interpret it as an argument, some parts of the text may have to be deleted, while other parts, especially missing assumptions that might be needed as premises or conclusions, have to be filled in. In order to begin this process, the critic needs to set up a key list of propositions that are explicitly stated in the argument, propositions that clearly represent statements the arguer has gone on record as claiming. But in order to proceed further with the process of analysis, in order to apply an argumentation scheme to this set of propositions, the critic typically needs to make up some conjecture, or hypothesis as it might be called, that accounts for the text having been put forward by the arguer in the way it was, in the context of dialog, including matters of what can reasonably be taken as common knowledge in relation to the text.

It looks like what the critic is doing at this point is to use abductive reasoning to set up a hypothesis that presumably offers the best explanation of the data given in the text of discourse. But in section 1, we drew a careful distinction between explanation and clarification, and added the qualification that interpretation of a text of discourse of the kind needed for critical argumentation is different from either of these other speech acts. Following along these lines, we could analyze what the critic is doing by saying that she is setting up an interpretation of the text of discourse as a hypothesis, or perhaps considering several competing interpretations side by side, and picking one of these interpretations as a starting point for the analysis, and ultimately for the evaluation of the given argument. If this account of what typically occurs during the process of critical argumentation is accurate, the form of reasoning central to it is not inference to the best explanation, but rather a kind of reasoning that could be called inference to the best interpretation. As the process of critical argumentation proceeds along these lines, the initial hypothesis, the chosen interpretation, functions as a template that can then be reapplied to the original text of discourse over and over again to draw out logical consequences that may be presumed to represent the arguer's implicit commitments.

Can this process of selecting a hypothesis as a best interpretation be categorized as a form of abductive reasoning? It may be difficult to so categorize it on the standard theory of abductive reasoning as inference to the best explanation, as presented by the Josephsons (1994) and Walton (2204). However, Gabbay and Woods (2005, p. 289) postulate a species of abductive reasoning they call interpretation abduction. On their theory (p. 290), the interpreter's target during the process of analyzing a text of discourse by bringing out implicit assumptions in it sometimes cannot be carried out by the direct method of asking the arguer what he means. In such a case the solution they propose is for the interpreter to conjecture a semantic hypothesis that takes the form of an interpretation of the given text

of discourse. Among the applications of interpretation abduction they cite is the determination of missing assumptions in enthymemes. Following along the lines of their proposal, it is suggested here that an interpretation of the text of discourse of the kind used for critical argumentation can be seen as based on a kind of abductive reasoning. It is a variant of abductive reasoning as inference to the best explanation that instead works as a process of inference to the best interpretation.

It is well to stress that several interpretations of the argumentation in the same text of discourse are often possible, and that a critic can very often come up with a simpler and a deeper analysis of the same given argument, both of which can be reasonable. Still, in some cases, some interpretations are better supported as more reasonable than others. Some can even commit the straw man fallacy. Thus even the best supported interpretation of the text is best seen as only a conjecture or hypothesis that is useful for purposes of critical argumentation, and evidence should count for or against it. Even the original arguer herself, if consulted, may be unsure about what she meant. Ambiguity is possible in any natural language text of discourse. It may even be an ambiguity unanticipated by the writer. If there is a choice between two very plausible interpretations, the critical evaluator of the argumentation in the discourse should point out the ambiguity. Then two argument reconstructions can be worked up and evaluated separately. The process of abductive interpretation will not always pick out exactly one interpretation as the only possible choice. But whatever interpretation is put forward as a hypothesis, it must stand or fall on the given textual evidence as data. Thus reproducible data are there, and thus drawing conclusions about an argument by interpretation abduction can be reasonable.

8   Textual interpretation as simulative

The reasoning used in interpreting a text of discourse is not only abductive, according to the theory of critical argumentation presented here. It is also simulative. Simulative reasoning is the kind of reasoning used when one agent reasons about the reasoning of another agent. The two agents are called the primary agent and the secondary agent. The secondary agent reasons about the reasoning of the primary agent. Of course, the secondary agent cannot directly enter into the mind of the primary agent. Simulative reasoning is indirect. It involves a process of what Collingwood (1946) called "reenactment". The secondary agent puts himself into the situation of the primary agent by imagining the primary agent solving a problem. Then the secondary agent imagines how he would proceed if he were confronted with that problem. According to Collingwood, historical explanation is accomplished by the process of reenactment, or simulative reasoning, as it is now called. It is interesting that Collingwood's leading example of this process of reen-

actment was a case of textual interpretation. Collingwood (1946, p. 283) gave the example of a contemporary historian reading the Theodosian code, the written edict of an ancient Roman emperor. According to Collingwood, what the historian must try to do is to imagine the problem the Roman emperor was trying to solve by writing the edict. She must try to imagine this problem as the emperor saw it. She must try to put herself in the emperor's situation. In Collingwood's view of historical explanation, the historian must simulate the thinking process the Roman emperor presumably went through to understand the reasons why he drafted the edict in the way he did. The historian must use simulative reasoning in reconstructing the thinking of the historical person whose traces have been left in historical documents, witness accounts, archaeological findings, and other historical evidence.

Simulation theory in psychology and cognitive science has been put forward as a theoretical model to explain animal and human behavior. Simulation theory has also had many critics, who find the notion of simulation mysterious and hard to define exactly, and who would prefer to have a theory of behavior that did not rest on a notion that is so hard to grasp. These critics have postulated an alternative to simulation theory called the theory-theory (explained below). The controversy about the two theories can be explained by citing a famous experiment. This experiment graphically suggested that the notion of simulative reasoning can be useful in understanding a kind of thinking that is found in animal behavior. In the experiment (Premack and Woodruff, 1978), a chimpanzee was shown a film of an actor trying unsuccessfully to reach for some bananas dangling overhead. Then the chimp was shown several pictures illustrating what the actor's next move could be. The picture selected by the chimp showed the actor moving some nearby crates underneath the bananas, and stacking the crates to provide a platform for reaching the bananas. The chimp's selection of this picture suggests the hypothesis that, based on its experience, the chimp was able to figure out the right method of solving the problem. One plausible explanation of the data found in the experiment is that the chimp arrived at the right solution to the problem by putting itself imaginatively into the situation of the actor shown in the picture. The chimp would, according to this hypothesis, imagine itself trying to reach for the bananas, and then, seeing the crates in the picture as the available means to this desirable end, imagine itself as dragging the crates over to the right area and stacking them up. The opposed hypothesis mentioned above is the so-called theory-theory. According to this hypothesis, the chimp accomplishes the feat of finding the right solution to the actor's problem through a process of goal-directed practical reasoning. The theory-theory approach suggests that all the chimp needs to grasp is a process of reasoning often called the Aristotelian practical syllogism. According to this theory, the only kind of thinking required is a linking of actions to goals. By this explanation of the data found in the experiment, the chimp has the goal of reaching the

bananas, or would find that a worthy goal. The chimp then reasons that in order to get the bananas, out of reach as they are depicted, the best available means would be to drag the crates under them, stack the crates, and then stand on them. According to the theory-theory explanation, no simulation is required because the chimp is "hard-wired" to understand how Aristotelian goal-directed practical reasoning works in a simple situation like that pictured in the bananas experiment. It doesn't have to simulate, or enter into the actor's mind imaginatively. It just sees the problem and reasons out a step by step solution.

Each theory has its problems. The theory-theory has to try to explain the data of this experiment, and others like it, without using any notion of simulation. The simulation theory has to try to explain what is meant by 'simulation' as a precise notion. A problem is that simulation could mean many things, and it could be construed more narrowly or inclusively. In psychology, an inclusive notion of simulation would seem to be most useful. According to this wide notion, one agent simulates the thinking of another agent by being able to imaginatively grasp the feelings and emotions of the other agent. Goldman (1995) adopted this wide notion when he proposed that 'simulation' means that one agent pretends to have the same beliefs and emotional reactions that another agent has. Goldman (1995, p. 189) described simulation as "pretending to have the same initial desires, beliefs, or other mental states that the attributor's background information suggests the agent has." This account of simulation was taken by Goldman (1995, p. 187) to include emotions like annoyance or "upsetness" that a person feels when delayed in traffic on the way to the airport. On the other hand, it is also possible to adopt a narrower account of what simulation is. According to this narrower account, simulation may not extend to gasping another person's feelings, but is restricted to trying to guess at what their feelings might be, based on data showing how that other person acts, what he says, and how he could be expected to draw conclusions, by practical reasoning for example. This narrower view of simulation makes simulation theory not all that far apart from the theory-theory. Thus the problem seems to turn on how to define 'simulation', an intuitively powerful idea that is hard to pin down in exact terms.

For various purposes, 'simulation' can be defined in a narrower or more inclusive way. For purposes of psychology, simulation might include one person's observing the other person's face, for example, and drawing inferences about that other person's emotional reactions based on his facial expressions. This notion of what constitutes simulation could be highly inclusive. But for purposes of informal logic, a typical case of concern would be one where a critic is trying to analyze and evaluate a specimen of argumentation found in a text of discourse, like a newspaper editorial or letter to the editor. To perform this task of analysis and evaluation of the writer's argumentation, the critic does not observe the writer's

face, or engage in immediate verbal confrontation with her. The simulation he performs is on her words. He has to try to judge fairly what she means, and how her argumentation can be interpreted. He may have to fill in missing premises, and make other interpretations of the text of discourse, based on what she may plausibly be taken to be saying. For this purpose, simulation is required. To fill in the missing parts charitably, he must try to insert himself into her way of thinking, and to sympathetically grasp her point of view in the argument. This skill is often called empathy. But in this kind of case, the simulation required is relatively narrow. The simulator needs to try to interpret what the writer's argument is, or may plausibly be taken to be, judging by what she has written in the text of discourse that he can see. He needs to grasp what her general position or viewpoint is, as expressed by this written discourse. Actually, in some ways this task is not unlike that of the chimp experiment. The chimp does not actually see the actor trying to reach the bananas. He only sees a picture of it.

The meaning of 'simulation' in which one person pretends to have the same beliefs and emotional reactions that another person has (Goldman, 1995, p. 189) may be appropriate for psychology, but it is too wide for the study of critical thinking (informal logic). Simulation of another agent's feelings, or other mental states is not necessary from the viewpoint of critical argumentation. Instead what is important is for the secondary agent to have the capability to simulate the reasoning of the primary agent, based on the reasoning capabilities that both have as agents. Both share a grasp of goal-directed practical reasoning as used in common kinds of situations both are familiar with, for example. One may not be able to simulate the goal or motive of the other. But the one agent can simulate how the other reasons, presuming the other has a certain goal that has been expressed verbally, to the conclusion to take certain steps to try to carry out that goal. Understandably, there is controversy about exactly what simulation is, and how widely or narrowly the notion should be construed for various purposes. But at least some fairly minimal notion of simulative reasoning is an essential element in developing a framework for interpretation that is useful for critical thinking (informal logic).

We now need to recall that in chapter 4, section 6, explanation was defined in terms of understanding of a kind that can be tested in an examination dialog. This kind of dialog is carried out by a questioner asking a series of questions to a respondent. The questioner's goal is to obtain information from the respondent, but to do this she often needs to try to get a clear and coherent account of what the respondent means, and to test the reliability of the information extracted from the respondent. To perform these functions, the information elicited is tested against the respondent's other statements or commitments, other known facts in the case, and other information thought to be true.

The formal analysis of examination dialog by Dunne, Doutre and Bench-Capon (2005) models the dialog structure of how this testing function of examination dialog essentially works. In the kind of case we are interested in here, an examination dialog begins with some text of discourse in natural language, like a paragraph written by an author who can be identified, that may have been taken from a book or news media source, and the title of the book may be known. The questioner in the dialogue is the argument analyst. She will identify the premises and the conclusion, look for missing premises, present the structure of the argument in the form of a diagram, and so forth.

The first part of the examination takes the form of a markup process whereby the analyst reconfigures the given text so that the argument contained in it is represented by a set of statements analyzed as premises or conclusions in a sequence of argumentation. Then there is a move to the level of criticism marking a shift to a critical discussion type of dialog. It is an odd type of critical discussion, because generally the author of the text is not present to take part in the dialog. His part in the dialog is fixed by the text. The critic takes the part of the questioner in the examination dialog, and conducts an exegesis of the respondent's text. The purpose of such an exegetical dialog is for the analyst to find the supposed meaning of the author. As suggested above, this process is based on abductive inference. There is a given body of data, namely the given text of discourse in the case, and the analyst is trying to give the best explanation of that text. She is trying to give an account of its meaning, or presumed meaning. A good exegesis should preserve the main line of argumentation that can be extracted from the text by an argument diagram, and it should fairly represent the arguer's position.

Examination dialog has two levels, and there is a dialectical shift in exegesis from the first level to the second. At the first level, the exegetical reconstruction of a text needs to be based on the reproducible textual evidence. At the second level, the questioner begins to plays the role of critic to probe into the argument and deal with ambiguity, vagueness, inconsistencies, and all the other problems of interpretation encountered. Formalizing the dialog structure of examination dialog is a promising direction for research in the project of developing automated tools for assisting in legal argumentation, and for seeing how automated examination dialog tools might be used in multi-agent systems.

9   Anticipating an arguer's future moves

The new abductive and simulative theory of textual intepretation can be extended to solve other problems that are vitally important to informal logic, rhetoric, artificial intelligence, and to the study of argumentation generally. What enables an

arguer to anticipate the arguments that will likely be used by another arguer, and to devise arguments, in advance, that will respond to those anticipated arguments? Evidently this ability is based on a kind of simulative reasoning. The one arguer can, it seems, simulate the viewpoint of the other somehow, and then extrapolate from that viewpoint to conjecture what kind of argument would suit that viewpoint. This ability appears to demand a temporary suspension of one's own viewpoint. It is something not everyone can do, and that some can do better than others. It seems to demand a kind of mental flexibility that is held to be a good quality for a critical thinker. But as those of us who have taught introduction to critical thinking courses know, some people find this kind of mental flexibility hard to grasp or exercise, while others are very good at it right away. It is a skill that requires imagination and empathy, as opposed to only being a skill of finding the strongest arguments to support a viewpoint. You have to enter imaginatively into the viewpoint of the other party, and then try to imagine how that other party would react to a particular argument.

The classic illustration of this ability is the performance of Socrates in the Platonic dialogs. Although Socrates claims not to have knowledge, he does have his own viewpoint and consistently argues for it, even though he does appear to be open to considering other views. But when he uses his technique of elenchtic questioning, he seems to have a pretty good idea, in advance of asking a question, of how the respondent is likely to answer it. It is this anticipation of how the respondent is likely to answer that enables him to lead the respondent into contradictions and other kinds of problematic outcomes. The basis of his elenchtic technique is this skill of being able to anticipate how the other party in a dialog will reply at future moves in the dialog. Anticipatory skills in argumentation may be comparable, in their use of simulative reasoning, to the kinds of strategic skills in structured games that have been studied by cognitive scientists.

Strategies in games also involve simulation of the thinking of another person. Gordon (1986) used the example of strategizing in chess to show how simulative reasoning is used when one player tries to anticipate the likely moves of the other player. According to Gordon (1986, p. 162), chess players report that they often visualize the chessboard from the opponent's point of view. In a simulation of the kind used in chess strategy, the chess player imagines herself as playing from the opponent's position. This reversal of positions involves a reversal of strategy. You now see yourself in the opponent's position as attacking your own position. The chess strategist is not trying to duplicate the actual beliefs or feelings of the other player. The simulation is more narrow than that. It is only a simulation of the logical thinking of the other player. The chess strategist is trying to simulate how she herself would develop strategies for planning out future moves if she were in the opponent's position. Simulative anticipation has not been studied, or paid much

attention to, in argumentation theory. But it would appear to be centrally important in several respects.

This simulative ability to anticipate an arguer's future moves is the basis of understanding how argumentation strategies work. This basis is, in turn, fundamental not only to rhetoric and persuasion, but also to understanding informal fallacies. Consider the example of fallacies of irrelevance, an important category of informal fallacies. The concept of relevance is also fundamental to evidence law. In court, one attorney might object that the argumentation of the opposing attorney is irrelevant. The opposing attorney might plead to the judge, "If your honor will give me a little latitude, I can show where this line of questioning is leading." How should the judge rule in such a case? What method should the judge use to arrive at a reasonable ruling in such a case? Evidently the answer is that the judge has to perform a kind of act of mental anticipation to try to estimate where the attorney's argument is leading, or is likely to lead. But of course, that is not known. It can only be conjectured by a mental act of extrapolating the line of argumentation forward, from the given point it has reached in the trial so far, to some conclusion it might arrive at in the future argumentation in the trial. In short, the judgment of relevance requires an anticipation of an arguer's future moves.

Similar problems have been encountered in designing web pages that might be used by different groups of people. For example a web site designed to give information about university courses and facilities might need to be directed to different groups of students and members of the community who might use that web page for various purposes. Some recent research (Bolchini and Paolini, 2006) has developed an interactive dialog model specifically tailored for such multichannel applications. One of the needs is to use the design model for brainstorming processes in which ideas are generated and discussed among stakeholders, developers and potential users of a web site. Dialog design is needed to define the structure of the kind of conversation that the designers anticipate will take place when the website is activated.

The problem of how argument anticipation works can be solved by using the basic structure of dialog in Hamblin's formal theory (1970; 1971). In a dialog, the proponent makes the first move, the respondent makes the next move, and then the dialog continues according to the rules, producing an orderly sequence of moves. As noted in chapter 5, section 3, each member in the sequence is defined by Hamblin (1971, p. 130) as a triple $\langle n, p, l \rangle$. n is the length of the dialog, defined as the number of moves made, p is a participant, and l is a locution. A small example of a dialog with three moves can be used to illustrate how Hamblin's formalism works.

$$\langle 0, P_0, L_4 \rangle, \langle 1, P_1, L_3 \rangle, \langle 2, P_0, L_2 \rangle$$

At move zero, P_0 begins the dialog by putting forward a locution of type 4. At move 1, P_1 replies by putting forward a locution of type 3. At move 2, P_0 replies with a locution of type 2. In Hamblin's formalism, a dialog of this sort always begins with an opening move (the zero move) and it ends with a closing move where the dialog is terminated. The sequence of moves between the opening and closing moves must follow the procedural rules in every dialog, meaning the locution rules, the dialog rules, the commitment rules and the success rules. Hamblin wanted to use formal dialog structures as part of a practical method for analyzing logical fallacies, which meant that the main task was taken to begin with the text of discourse that supposedly contains some sort of argument that needs to be identified, analyzed, and evaluated as fallacious or not. This task requires looking backwards to examine what one's partner in the dialog previously said, by examining the moves he previously made in the dialog and attempting to determine what his commitments are, based on those moves.

But what about looking forward, where one participant tries to look ahead to anticipate what the other participant might do, what move he might be likely to make, at some future move in the dialog? For example, looking at the sequence above, could P_1 at move 1 try to predict whether or not P_0 will put forward locution L_2 or not?

One of the most important tools in exercising skills of anticipating another party's line of argument in a dialog is the argumentation scheme with its matching set of critical questions. The critical questions represent the most common kinds of doubts and objections that an arguer who is responding to one's own previous arguments might have. As noted in chapter 5, the procataleptic function of an argumentation scheme with its matching set of critical questions is to strengthen an argument by dealing with possible criticisms, objections and counter-arguments before they are even raised in a dialog. Of course, that tool by itself is not sufficient, because there are all kinds of counter-arguments to one's previous argument that might be plausible, over and above the standard set of critical questions.

The use of schemes needs to be supplemented by, and indeed seen as part of, the new abductive and simulative theory of argument interpretation. One of the main tools used in this theory is the argument diagramming tool of the kind exemplified by *Araucaria* and Carneades. Other tools are the dialog systems cited as useful for different purposes in this book, systems like CB, ASD, CBV and CBE. These systems offer resources for helping an argumentation strategist look backwards to the data provided by the argumentation reached at a given point. He must examine each of the arguer's previous moves in the dialog, and most importantly, he must determine what the arguer's commitments are, with respect to the possibility of his making the next move. Then he must look forward, and by simulating the reasoning of the proponent, he must construct hypotheses about where the argument may plausibly lead, based on that given data. The method is essentially

abductive and simulative. The argumentation strategist has to extrapolate ahead, based on the data already known about how the argumentation has progressed to the given point in the case. The skill is one of guessing where the line of argumentation will plausibly lead, given what is known about where it is now.

The conclusion is that the skill of looking forward in a dialog, to try to make plausible conjectures about what one's speech partner's future moves are likely to be, depends on one's capability for looking backward in the dialog. The evidence for judging where the dialog is likely to go in the future is to be found in the moves the other party made in the past. Skills of argument anticipation are therefore directly dependent on skills of argument interpretation.

10   The problem of diffuse dialog

At first sight, critical interpretation of a text of discourse does not seem to be dialectical. The critic does not actually engage in an active dialog with the writer (or speaker) of the given text of discourse. Instead, the text is simply presented. The proponent of the argumentation presented in the text (normally, or typically) is not actually present to defend his viewpoint in verbal exchanges with the critic. So you could say that it is not a real dialog. The one party is passive. The same is true of many cases in rhetoric. A speaker gives a televised political speech. The audience reacts to it. They may be persuaded by it or not. Some may like it and some may reject the arguments in it. But the speech is just there as a given text of discourse. Critics may evaluate the argumentation in the speech without the speaker getting a chance to criticize the critic.

The problem in this kind of case is that only the one party in the dialog, the proponent, is actually speaking, while the respondent is only hypothesized or conjectured. Indeed, in the case of a TV ad or a newspaper editorial, the projected audience for readership consists not of one individual, but of many the proponent has not even met. He cannot even know in advance who these individual respondents are, and can only conjecture what their commitments are, based on his own comparable commitments, and the ones he attributes to groups of people to whom these respondents belong. The framework of this kind of argumentation could be called that of diffuse dialog, because the respondent can only imagine the commitments of the audience and the critical questions that they might be likely to ask. The respondent is not one specific individual, but rather a category, or perhaps even several categories with overlapping interests and commitments. The study of how dialog theory can be extended to apply to diffuse dialogs is a research project for the future – a next big step for critical argumentation.

Diffuse dialog can be described as a communication artifact, and it works in a somewhat different way from Hamblin's model in certain respects. But it can still be studied and formalized using the formal dialog models advanced in this book. The reason this is so is that the proponent can look ahead beyond the given sequence of moves in the dialog so far, through a series of imagined turns in which the respondent is hypothesized as plausibly making certain kinds of moves as those turns come up. As shown in chapter 3, section 4, argument-based logics model argumentation as a tree superimposed on another tree (Prakken, 2007, p.6). First there is a tree, for example an argument tree as represented in an *Araucaria* diagram, in which propositions are linked to each other by inferences warranted by argumentation schemes. Second, the dialectical status of arguments is tracked in a secondary tree that maps possible ways to support or attack a given argument. In this framework, the one party in the dialog projects possible dialog sequences forward as a structure in which alternative future moves of one's opponent in the dialog are conjectured. What is the evidence on which such a conjecture needs to be based? As concluded in the last section, the evidence should be based on what the proponent knows about the past performances of the respondent, including common knowledge and the arguer's commitments. On this theory, argument evaluation and argument strategy are both based on the same underlying dialog structure. A consequence of the theory is that dialectic, which is centrally concerned with argument evaluation, and rhetoric, which is centrally concerned with argument invention and strategy, are not as far apart as tradition assumes. The new rhetoric will need to be based on the new dialectic as its underlying structure.

According to the abductive and simulative theory of textual interpretation presented in this chapter, textual interpretation for critical argumentation is best viewed as dialectical. It is dialectical because the critic must enter into the viewpoint of the proponent of the argumentation in order to explain what she presumably means. The critic must use the data provided by the text and by this simulation of the proponent's thinking propose a hypothesis, or a set of hypotheses, that explain that meaning. The critic's interpretation is best seen as an assumption that is tentative, and is itself open to critical questioning. A different observer may see the original argument in a different light, and may see a new way of interpreting it that the original critic overlooked. The critic should see his proposed interpretation as a hypothesis that is open to discussion and to the possibility of being tested against alternative interpretations. The criteria for evaluating abductive arguments, as shown above, indicate that pragmatic factors of the context of dialog are an important part of the evidence to be considered.

In critical argumentation, there is a special purpose for interpreting a text of discourse, and special methods need to be used. The study of fallacies has shown that the purpose of the discourse in a given case can be very important evidence in

deciding how to evaluate the argumentation in that case. The very same argument, or local sequence of reasoning from premises to a conclusion, could be evaluated quite differently if used in different kinds of discourse. For example, consider an argument used in a commercial ad to promote a product. Then consider the same argument used in a *Consumer's Report* article to evaluate the product by testing it against competing products. The very same argument could be used quite appropriately in the commercial ad, yet be biased and subject to severe criticism if used in the *Consumer's Report* article. So context of use is important. The purpose of the discourse is important. This relativity of argumentation to context of use in different types of dialog has been shown, over and over again, in studying the various informal fallacies. The purpose of a discourse is not only important for evaluating argumentation, however. It is also important for interpreting argumentation.

Bibliography

Aakhus, Mark 2006. "The Act and Activity of Proposing in Deliberation". In *Engaging Argument: Selected Papers from the 2005 NCA/AFA Summer Conference on Argumentation*, National Communication Association, 402–408.

Acock, Malcom 1985. *Informal Logic Examples and Exercises.* Belmont: Wadsworth.

Alexy, Robert 1989. *A Theory of Legal Argumentation.* Oxford: Clarendon Press.

Aristotle 1939. *Topics* (trans. E. S. Forster). Cambridge: Harvard University Press.

Aristotle 1928. *On Sophistical Refutations.* Cambridge: Harvard University Press.

Aristotle 1984. *Nicomachean* Ethics (trans. W. D. Ross, revised by J. O. Urmson). In Jonathan Barnes (ed.), *The Complete Works of Aristotle.* Princeton: Princeton University Press.

Atkinson, Katie, Bench-Capon, Trevor and McBurney, Peter 2004. "Computational Representation of Persuasive Argument". Technical Report ULCS-04–006, Department of Computer Science, University of Liverpool, UK.

Atkinson, Katie, Bench-Capon, Trevor and McBurney, Peter 2004. "PARMENIDES: Facilitating Democratic Debate". In R. Traunmuller (ed.), *Electronic Government, Lecture Notes in Computer Science (LNCS), 3183. Third International Conference on eGovernment (EGOV 2004), DEXA 2004,* Zaragoza, Spain.

Atkinson, Katie, Bench-Capon, Trevor and McBurney, Peter 2005. "Agent Decision Making Using Argumentation About Actions". Technical Report ULCS-05–006, University of Liverpool, Computer Science Department.

Atkinson, Katie, Bench-Capon, Trevor and McBurney, Peter 2006. "Computational Representation of Practical Argument". *Synthese* 152: 157–206.

Audi, Robert 1989. *Practical Reasoning.* London: Routledge.

Barnden, John A. 1995. "Simulative Reasoning, Common-sense Psychology, and Artificial Intelligence". In Martin Davies and Tony Stone (eds.), *Mental Simulation*, Oxford: Blackwell, 247–273.

Barnes, Jonathan 1980. "Aristotle and the Methods of Ethics". *Revue Internationale de Philosophie* 34: 590–511.

Barth, Else M. and Krabbe, Erik C. W. 1982. *From Axiom to Dialogue.* New York: De Gruyter.

Bench-Capon, Trevor J. M. 1995. "Argument in Artificial Intelligence and Law". In J. C. Hage, T. J. M. Bench-Capon, M. J. Cohen and H J. van den Herik (eds.), *Legal Knowledge Based Systems: JURIX 95, The Foundation for Legal Knowledge Systems,* Lelystad: Koninklijke Vermande BV, 5–14.

Bench-Capon, Trevor J. M. and Staniford, G. 1995. "PLAID – Proactive Legal Assistance". In *Proceedings of the Fifth International Conference on AI and Law, University of Maryland.* New York: ACM Press, 81–88.

Bench-Capon, Trevor J. M. 2002. "Agreeing to Differ: Modelling Persuasive Dialogue between Parties with Different Values". *Informal Logic* 22: 231–245. Version available (2003) on the

web page of the Department of Computer Science at the University of Liverpool: http://www.csc.liv.ac.uk/~tbc/projects/ifl.pdf

Bench-Capon, Trevor J. M. 2003. "Persuasion in Practical Argument Using Value-based Argumentation Frameworks". *Journal of Logic and Computation* 13: 429–448.

Bitzer, Lloyd F. 1959. "Aristotle's Enthymeme Revisited". *The Quarterly Journal of Speech* XLV: 399–408.

Black, Max 1946. *Critical Thinking*. New York: Prentice Hall.

Boger, George 2005. "Subordinating Truth: Is Acceptability Acceptable?". *Argumentation* 19: 187–238.

Bolchini, Davide and Paolini, Paolo 2006. "Interactive Dialogue Model: A Design Technique for Multichannel Applications". *IEEE Transactions on Multimedia* 8: 529–541.

Bolton, Robert 1990. "The Epistemological Basis of Aristotelian Dialectic". In Daniel Devereux and Pierre Pellegrin (eds.), *Biologie, Logique et Metaphysique chez Aristote*, Paris: Editions du Centre de la Recherche Scientifique, 185–236.

Bratman, Michael 1897. *Intentions, Plans, and Practical Reason*. Cambridge: Harvard University Press.

Bratman, Michael E., Israel, David J. and. Pollack, Martha E. 1988. "Plans and Resource-bounded Practical Reasoning". *Computational Intelligence* 4: 349–355.

Brinton, Alan 1986. "Ethotic Argument". *History of Philosophy Quarterly* 3. 245–257.

Bromberger, Sylvain 1966. "Why-Questions". In Robert G. Colodny (ed.), *Mind and Cosmos*, Pittsburgh: University of Pittsburgh Press, 86–111.

Burke, Michael 1985. "Unstated Premises". *Informal Logic* 7: 107–118.

Burnyeat, Myles F. 1994. "Enthyme: Aristotle on the Logic of Persuasion". In David J. Furley and Alexander Nehemas (eds.), *Aristotle's Rhetoric: Philosophical Essays*, Princeton, N. J.: Princeton University Press, 3–55.

Carberry, Sandra 1990. *Plan Recognition in Natural Language Dialogue*. Cambridge: MIT Press.

Carbogim, Daniela V., Robertson, David and Lee, John 2000. "Argument-Based Applications to Knowledge Engineering". *The Knowledge Engineering Review* 15: 119–149.

Carlson, Lauri 1983. *Dialogue Games: An Approach to Discourse Analysis*. Dordrecht: Reidel.

Castelfranchi, Cristiano 1995. "Commitments: From Individual Intentions to Groups and Organizations". In Victor Lesser (ed.), *Proceedings: First International Conference on Multi-Agent Systems*, Menlo Park: AAAI Press.

Cawsey, Alison 1992. *Explanation and Interaction: The Computer Generation of Explanatory Dialogue*. Cambridge: MIT Press.

Chandrasekaran, B. 1986. "Generic Tasks in Knowledge Based Reasoning". *IEEE Expert* 1: 23–30.

Chu-Carroll, Jennifer and Carberry, Sandra 1995. "Conflict Detection and Resolution in Collaborative Planning". In *Intelligent Agents II, Agent Theories, Architectures and Languages: IJCAI '95 Workshop*, Berlin: Springer-Verlag, 111–126.

Clarke, David S. Jr., 1985. *Practical Inferences*. London: Routledge.

Coffa, J. A. 1974. "Hempel's Ambiguity". *Synthese* 28: 141–163.

Cohen, P. R. and Levesque, H. J. 1995. "Communicative Actions for Artificial Agents". In *Proceedings of the First International Conference on Multi-Agent Systems*, 65–72.

Collingwood, Robin G. 1946. *The Idea of History*. Oxford: Clarendon Press.

Copi, Irving M. and Cohen, Carl 1990. *Introduction to Logic*. New York: Macmillan.

Craig, Robert T. 1999. "Communication Theory as a Field". *Communication Theory*, to appear.

Cushing, Steven 1994. *Fatal Words: Communication Clashes and Aircraft Crashes*, Chicago: University of Chicago Press.

Dascal, Marcelo 2003. *Interpretation and Understanding.* Amsterdam: John Benjamins.

Dascal, Marcelo and Gross, Alan G. 1999. "The Marriage of Pragmatics and Rhetoric". *Philosophy and Rhetoric* 32: 107–130.

Dascal, Marcelo, Eemeren, Frans H. van, Rigotti, Eddo, Stati, Sorin and Rocci, Andrea 2005. "Dialogue, Argument, Controversy". *Studies in Communication Sciences, Special Issue on Argumentation in Dialogic Interaction*: 1–20.

Devereux, Daniel 1990. "Comments on Robert Bolton's The Epistemological Basis of Aristotelian Dialectic". In Daniel Devereux and Pierre Pellegrin (eds.), *Biologie, Logique et Mtephysique chez Aristote*, Paris: Editions du Centre National de las Recherche Scientifique, 263–286.

Dignum, Frank and Greaves, Mark 2000. "Issues in Agent Communication: An Introduction". In *Issues in Agent Communication*, Berlin, Springer-Verlag, 1–16.

Dove, Kenley R. 1970. "Hegel's Phenomenological Method". *The Review of Metaphysics* 23: 622–641.

Dray, William 1964. *Philosophy of History.* Englewood Cliffs: Prentice-Hall.

Dray, William 1995. *History as Re-enactment: R. G. Collingwood's Idea of History*, Oxford: Oxford University Press.

Dung, P. M. 1995. "On the Acceptability of Arguments and Its Fundamental Role in Nonmonotonic Reasoning, Logic Programming and n-person Games". *Artificial Intelligence* 77: 321–357.

Dunne, Paul E. and Bench-Capon, Trevor J. M. 2006. *Computational Models of Argument: Proceedings of COMMA 2006.* Amsterdam: IOS Press..

Dunne, Paul E., Doutre, Sivlie, and Bench-Capon, Trevor, 2005. Discovering Inconsistency through Examination Dialogues, *Proceedings IJCAI-05 (International Joint Conferences on Artificial Intelligence)*, Edinburgh, 1560–1561. Available at: http://ijcai.org/search.php

Ennis, Robert H. 1982. "Identifying Implicit Assumptions". *Synthese* 51: 61–86.

Evans, J. D.G. 1977. *Aristotle's Concept of Dialectic.* London: Cambridge University Press.

Farley, Arthur M. and Freeman, Kathleen 1995. "Burden of Proof in Legal Argumentation". In *5th International Conference on Artificial Intelligence and Law*, College Park, Maryland, 156–164: http://portal.acm.org/citation.cfm?id=222227

Farrell, Thomas B. 1993. *Norms of Rhetorical Culture.* New Haven: Yale University Press.

Farrell, Thomas B. 2000. "Aristotle's Enthymeme as Tacit Reference". In Alan G. Gross and Arthur E. Walzer (eds.), *Rereading Aristotle's Rhetoric,* Carbondale: Southern Illinois University Press, 93–106.

Federal Rules of Evidence (FRE) 1997, A Hypertext Publication, Legal Information Institute, http:www.law.cornell.edu/rules/fre.

Felscher, Walter 1985. "Dialogues, Strategies, and Intuitionistic Provability". *Annals of Pure and Applied Logic* 28: 217–254.

Ferber, Jacques 1999. *Multi-Agent Systems: An Introduction to Distributed Artificial Intelligence.* Harlow, England: Addison-Wesley.

Feteris, Eveline T. 1999. *Fundamentals of Legal Argumentation.* Dordrect: Foris.

Finin, Tim, Fritzon, Richard, McKay, Don and McEntire, Robin 1994. "KQML as an Agent Communication Language". In *Proceeding of the Third International Conference on Information and Knowledge Management*, New York: ACM Press, 456–463.

Finocchiaro, Maurice 1975. "Cause, Explanation and Understanding in Science: Galileo's *Case*". *Review of Metaphysics* 29: 117–128.

Finocchiaro, Maurice 1980. "Scientific Discoveries as Growth of Understanding: The Case of Newton's Gravitation". In Thomas Nickles (ed.), *Scientific Discovery, Logic, and Rationality*, Dordrecht: Reidel, 235–255.

FIPA Agent Communication Language Specification. http://www.fipa.org/

Fischer, David Hackett 1970. *Historians' Fallacies.* New York: Harper and Row.

Flowers, M. and Dyer, M. E. 1984. "Really Arguing with your Computer". In *Proceedings of the National Computer Conference*, 653–659.

Franklin, Stan and Graesser, Art 1996. "Is It an Agent, or Just a Program? A Taxonomy for Autonomous Agents". In Jorg P. Muller, Michael J. Wooldridge and Nicholas R. Jennings (eds.), *Intelligent Agents III: Agent Theories, Architectures and Languages*, Berlin: Springer, 21–35.

Frapolli, Maria J. 1992. "Identity, Necessity and *a Prioricity*: The Fallacy of Equivocation". *History and Philosophy of Logic* 13: 91–109.

Freeman, James B. 1991. *Dialectics and the Macrostructure of Arguments.* Berlin: Foris.

Freeman, James B. 1995. "The Appeal to Popularity and Presumption by Common Knowledge". In Hans V. Hansen and Robert C. Pinto (eds.), *Fallacies: Classical and Contemporary Readings*, University Park, Pa.: The Pennsylvania State University Press, 263–273.

Freeman, James B. 1988. *Thinking Logically.* Englewood Cliffs: Prentice Hall.

Freeman, James B. 2005. *Acceptable Premises: An Epistemic Approach to an Informal Logic Problem.* New York: Cambridge University Press.

Friedrich, Carl J. (ed.) 1953. *The Philosophy of Hegel.* New York: Random House.

Gabbay, Dov M. and Woods, John 2005. *The Reach of Abduction: Insight and Trial.* Amsterdam: Elsevier.

Gagarin, Michael 1994. "Probability and Persuasion: Plato and Early Greek Rhetoric". In Ian Worthington (ed.), *Persuasion: Greek Rhetoric in Action*, London: Routledge, 46–68.

Girle, Rod, Hitchcock, David, McBurney, Peter and Verheij, Bart 2003. "Decision Support for Practical Reasoning: A Theoretical and Computational Perspective". In Chris Reed and Timothy J. Norman (eds.), *Argumentation Machines: New Frontiers in Argument and Computation*, Dordrecht: Kluwer, 58–84.

Goldman, Alvin 1970. *A Theory of Human Action.* Englewood Cliffs. Prentice-Hall.

Goldman, Alvin 1995. "Empathy Mind and Morals". In Martin Davies and Tony Stone (eds.), *Mental Simulation: Evaluations and Applications*, Oxford: Blackwell, 185–208.

Gordon, Robert M. 1986. "Folk Psychology as Simulation". *Mind and Language* 1: 158–171.

Gordon, Thomas F. 1995. *The Pleadings Game. An Artificial Intelligence Model of Procedural Justice.* Dordrecht: Kluwer Academic Publishers.

Gordon, Thomas F. and Karacapilidis, Nikos 1997. "The Zeno Argumentation Framework". In *Proceedings of the Sixth International Conference on Artificial Intelligence and Law, Melbourne*, Australia, 10–18.

Gordon, Thomas F. and Walton, Douglas 2006. "The Carneades Argumentation Framework: Using Presumptions and Exceptions to Model Critical Questions", *Proceedings of 6th CMNA (Computational Models of Natural Argument) Workshop*, ECAI (European Conference on Artificial Intelligence), Riva del Garda, Italy, August 28 – September 1, Trento, Italy, University of Trento, 5–13.

Gough, James and Tindale, Christopher 1985. "Hidden or Missing Premises". *Informal Logic* 7: 99–106.

Govier, Trudy 1982. "What's Wrong with Slippery Slope Arguments?". *Canadian Journal of Philosophy* 12: 303–316.

Govier, Trudy 1987. *Problems in Argument Analysis and Evaluation.* Dordrecht: Foris.

Govier, Trudy 1992. *A Practical Study of Argument*, 3rd ed. Belmont: Wadsworth.

Greco, Sara 2005. "Dascal on Interpretation and Understanding". *Studies in Communication Sciences* 5: 217–230.

Greenwood, Katie, Bench-Capon, Trevor J. M. and McBurney, Peter 2003. "Towards a Computational Account of Persuasion in Law". In *Proceedings of the Ninth International Conference on Artificial Intelligence and Law,* New York: ACM Press, 22–31.

Grice, H. Paul 1975. "Logic and Conversation". In Donald Davidson and Gilbert Harman (eds.), *The Logic of Grammar,* Encino, California: Dickenson, 64–75.

Gross, Alan G. 2000. "What Aristotle Meant by Rhetoric". In Alan G. Gross and Arthur E. Walzer (eds.), *Rereading Aristotle's Rhetoric*, Carbondale: Southern Illinois University Press, 24–37.

Hage, Jaap C., Leenes, Ronald and Lodder, Arno R. 1994. "Hard Cases: A Procedural Approach". *Artificial Intelligence and Law* 2: 113–167.

Hamblin, Charles L. 1970. *Fallacies.* London: Methuen.

Hamblin, Charles L. 1971. "Mathematical Models of Dialogue". *Theoria* 37: 130–155.

Hamblin, Charles L. 1987. *Imperatives.* New York: Blackwell.

Hamilton, Sir William 1861. *Discussions on Philosophy and Literature.* New York: Harper and Brothers.

Hamilton, Sir William 1874. *Lectures on Logic.* Edinburgh: William Blackwood and Sons.

Hart, H. L. A 1957–58. "Positivism and the Separation of Law and Morals". *Harvard Law Review* 71: 593–629.

Hastings, Arthur C. 1963. *A Reformulation of the Modes of Reasoning in Argumentation*, Evanston, Illinois: Ph.D. Dissertation.

Hauser, Marc D. 1996. *The Evolution of Communication.* Cambridge: MIT Press.

Hempel, Carl G. 1965. *Aspects of Scientific Explanation.* New York: The Free Press.

Hintikka, Jaakko 1979. "Information-Seeking Dialogues: A Model". *Erkenntnis* 38: 355–368.

Hintikka, Jaakko 1992. "The Interrogative Model of Inquiry as a General Theory of Argumentation". *Communication and Cognition* 25: 221–242.

Hintikka, Jaakko 1993. "Socratic Questioning, Logic and Rhetoric". *Revue Internationale de Philosophie* 1 (184): 5–30.

Hintikka, Jaakko 1995. "The Games of Logic and the Games of Inquiry". *Dialectica* 49: 229–249.

Hintikka, Jaakko and Hintikka, Merrill B. 1982. "Sherlock Holmes Confronts Modern Logic: Toward a Theory of Information-Seeking Through Questioning". In E. M. Barth and J. L. Martens (eds.), *Argumentation: Approaches to Theory Formation*, Amsterdam: Benjamins, 55–76.

Hitchcock, David 1985. "Enthymematic Arguments". *Informal Logic* 7: 83–97.

Hitchcock, David, McBurney Peter and Parsons, Simon 2001. "A Framework for Deliberation Dialogues". In H. V. Hansen, C. W. Tindale, J. A. Blair and R. H. Johnson (eds.), *Argument and Its Applications: Proceedings of the Fourth Biennial Conference of the Ontario Society for the Study f Argumentation (OSSA 2001),* compact disk. Also available on Peter McBurney's web page at the University of Liverpool, Department of Computer Science: http://www.csc.liv.ac.uk/~peter/

Hohmann, Hanns 2000. "Rhetoric and Dialectic: Some Historical and Legal Perspectives". *Argumentation* 14: 223–234.

Houtlosser, Peter and van Eemeren, Frans 1998. "Rhetorical Ways of Managing Disagreement: Justifying Reconstructions of Confrontation". In James F. Klumpp (ed.), *Argument in a*

Time of Change: Proceedings of the Tenth NCA/AFA Conference on Argumentation, Annandale, Va.: National Communication Association, 56–62.

Huhns, Michael N. and Singh, Munindar P. 1998. "Cognitive Agents". *IEEE Internet Computing*, November-December: 87–89.

Hurley, Patrick J. 1991. *A Concise Introduction to Logic* 4th ed.. Belmont, California: Wadworth.

Hurley, Patrick J. 2000. *A Concise Introduction to Logic*, 7th ed.. Belmont, California: Wadworth.

Hurley, Patrick J. 2003. *A Concise Introduction to Logic*, 8th ed.. Belmont, California: Wadworth.

Irwin, Terence 1998. *Aristotle's First Principles*. Oxford: Clarendon Press.

Jacobs, Scott 2000. "Rhetoric and Dialectic from the Standpoint of Normative Pragmatics". *Argumentation* 14: 261–286.

Jennings, Nicholas R., Sycara, Katia and Woodridge, Michael 1998. "A Roadmap of Agent Research and Development". *Autonomous Agents and Multi-Agent Systems* 1: 7–38.

Jennings Nicholas R. and Wooldridge, Michael 1995. "Applying Agent Technology". *Applied Artificial Intelligence* 9: 357–369.

Johnson, Ralph H. 1981. "Charity Begins at Home". *Informal Logic Newsletter* 3: 4–9.

Johnson, Ralph H. 2000. *Manifest Rationality: A Pragmatic Theory of Argument,* Mahwah, N. J.: Erlbaum.

Johnson, Ralph, H. and Blair, J. Anthony 1983. *Logical Self-Defence*, 2nd ed.. Toronto: McGraw-Hill Ryerson.

Johnson, Ralph H. and Blair, J. Anthony 1985. "Informal Logic: The Past Five Years". *American Philosophical Quarterly* 22: 181–196.

Joseph, H. W. B. 1916. *An Introduction to Logic*. Oxford: Clarendon Press.

Josephson, John R. and Josephson, Susan G. 1994. *Abductive Inference: Computation, Philosophy, Technology*. New York: Cambridge University Press.

Kahn, Charles M. 1996. *Plato and the Socratic Dialogue.* Cambridge: Cambridge University Press.

Kant, Immanuel 1961. *Critique of Pure Reason*, trans. Norman Kemp Smith. London: Macmillan & Co.

Kapp, Ernst 1942. *Greek Foundations of Traditional Thought*. New York: Columbia University Press.

Karunatillake, Nishan C., Jennings, Nicholas R., Rahwan, Iyad and Norman, Timothy J. (eds.) 2005. *Proceedings of the Fourth International Joint Conference on Autonomous agents and Multiagent Systems*. New York: ACP Press.

Kestler, Jeffrey L. 1982. *Questioning Techniques and Tactics*. New York: McGraw-Hill.

Kienpointner, Manfred 1987. "Towards a Typology of Argumentation Schemes". In Frans H. van Eemeren, Rob Grootendorst, J. Anthony Blair and Charles A. Willard (eds.), *Argumentation: Across the Lines of Discipline*, Dordrecht: Foris, 275–287.

Kienpointner, Manfred 1992. *Alltagslogik: Struktur und Funktion von Argumentationsmustern*. Stuttgart: Fromman-Holzboog.

Kienpointner, Manfred 1993. "The Empirical Relevance of Perelman's New Rhetoric". *Argumentation* 7: 419–437.

Kienpointner, Manfred 1997. "On the Art of Finding Arguments: What Ancient and Modern Masters of Invention Have to Tell Us About the *Ars Inveniendi*". *Argumentation* 11: 225–236.

Kimura, D.1993. *Neuromotor Mechanisms in Human Communication*. Oxford: Oxford University Press.

Kirkham, Richard 1992. *Theories of Truth: A Critical Introduction*. Cambridge, Mass.: MIT Press.

Kitcher, Philip 1988. "Explanatory Unification". In J. C. Pitt (ed.), *Theories of Explanation*, New York: Oxford University Press, 167–187.

Kneale William and Kneale, Martha 1962. *The Development of Logic*. Oxford: Oxford Clarendon Press.

Krabbe, Erik C. W. 1992. "So What? Profiles for Relevance Criticism in Persuasion Dialogues". *Argumentation* 6: 271–283.

Krabbe, Erik C. W. 1995. "Appeal to Ignorance". In H.V. Hansen and R.C. Pinto (eds.), *Fallacies: Classical and Contemporary Readings*, University Park: Pennsylvania State University Press, 251–264.

Krabbe, Erik C. W. 1998. "Dialogical Logic". In Edward Craig (ed.), *Routledge Encyclopedia of Philosophy*, London: Routledge, 59–62.

Krabbe, Erik C. W. 1999. "Profiles of Dialogue". In J. Gerbrandy, M. Marx, M. de Rijke and Y. Venema (eds.), *JFAK: Essays Dedicated to Johan van Benthem on the Occasion of his 50th Birthday*, Amsterdam: Amsterdam University Press, 25–36.

Krabbe, Erik C. W. 2000. "Meeting in the House of Callias: Rhetoric and Dialectic". *Argumentation* 14: 205–217.

Krabbe, Erik C. W. 2003. "Metadialogues". In F.H. van Eemeren, J.A. Blair, C.A. Willard and A. F. Snoek Henkemans (eds.), *Anyone Who Has a View: Theoretical Contributions to the Study of Argumentation*, Dordrecht: Kluwer, 83–90.

Krabbe, Erik C. W. "Dialogue Logic". In Dov M. Grabbay and John Woods (eds), *Handbook of the History of Logic*, vol. 1, Amsterdam, North-Holland, 2006, 665–704.

Kripke, Saul 1965. "Semantical Analysis of Intuitionistic Logic I". In J. N. Crossley and M. Dummett (eds.), *Formal Systems and Recursive Functions*, Amsterdam: North-Holland, 92–113.

Kripke, Saul 1975. "Outline of a Theory of Truth". *Journal of Philosophy* 72: 690–716.

Krothapalli, N.K.C. and Desmurkh, A.V. 1999. "Design of Negotiation Protocols for Multi-Agent Manufacturing Systems". *International Journal of Production Research* 37: 1601–1624.

Kuipers, Theo A. (ed.) 1987. What is Closer-to-the-truth? A Parade of Approaches to Truthlikeness. Poznan Studies in the Philosophy of the Sciences and the Humanities, Volume 10, Amsterdam: Rodopi.

Labrou, Yannis, Finin, Tim and Peng, Yun 1999. "Agent Communication Languages: The Current Landscape". *IEEE Intelligent Systems and Their Applications* 14: 45–52.

Leff, Michael 1983. "The Topics of Argumentative Invention in Latin Rhetorical Theory from Cicero to Boethius". *Rhetorica* 1: 23–44.

Leff, Michael 1993. "The Uses of Aristotle's Rhetoric in Contemporary American Scholarship". *Argumentation* 7: 313–327.

Leff, Michael 2000. "Rhetoric and Dialectic in the Twenty-First Century". *Argumentation* 14: 241–254.

Leibniz, Gottfried Wilhelm 2006. *The Art of Controversies*, edited by Marcelo Dascal, Quintín Racionero and Adelino Cardoso. Dordercht: Springer.

Lester, James C.and Porter, Bruce W. 1997. "Developing and Empirically Evaluating Robust Explanation Generators: The KNIGHT Experiments". *Computational Linguistics*, 23: 65–101.

Lodder, Arno R. 1998. *Dialaw: On Legal Justification and Dialog Games*. Ph. D. Thesis, Maastricht University.

Lodder, Arno R. 1999. *Dialaw: On Legal Justification and Dialogical Models of Argumentation*. Dordrecht: Kluwer.

Long, A. and Sedley, D. N. 1987. *The Hellenistic Philosophers: Translations of the Principal Sources with Philosophical Commentary*, vol. 1. Cambridge: Cambridge University Press.

Lorenzen, Paul 1969. *Normative Logic and Ethics*. Mannheim: Bibliographsiches Institut.

Lorenzen, Paul and Lorenz, Kuno 1978. *Dialogische Logik*. Darmstadt: Wissenschftliche Buchgesellschaft.

Mackenzie, Jim D. 1979. "Question-Begging in Non-Cumulative Systems". *Journal of Philosophical Logic* 8: 117–133.

Mackenzie, Jim 1981. "The Dialectics of Logic". *Logique et Analyse* 94: 159–177.

Mackenzie, Jim D. 1984. "Begging the Question in Dialogue". *Australasian Journal of Philosophy* 62: 174–181.

Mackenzie, Jim D. 1988. "Distinguo: The Response to Equivocation". *Argumentation* 2: 465–482.

Mackenzie, Jim 1990. "Four Dialogue Systems". *Studia Logica* 49: 567–583.

McAdon, Brian 2001. "Rhetoric is a Counterpart of Dialectic". *Philosophy and Rhetoric* 34: 113–149.

McBurney, James H. 1936. "The Place of the Enthymeme in Rhetorical Theory". *Speech Monographs* 3: 49–74.

McBurney, Peter and Parsons, Simon 2002. "A Geometric Semantics for Dialogue Game Protocols for Autonomous Agent Interaction". *Electronic Notes in Theoretical Computer Science* 52 (2): http://www.elsevier.nl/locate/entcs/volume52.html18page

Miller, David 1974. "Popper's Qualitative Theory of Verisimilitude". *The British Journal for the Philosophy of Science* 25:166–177.

Miller, Geoffrey P. 1990. "Pragmatics and the Maxims of Interpretation". *Wisconsin Law Review* 20: 1179–1227.

Moore, Johanna D. 1991. "A Reacting Approach to Explanation: Taking the User's Feedback into Account". In C. L. Paris, W. R. Swartout and W. C. Mann (eds.), *Natural Language Generation in Artificial Intelligence and Computational Linguistics*, Dordrecht: Kluwer, 3–48.

Moore, Johanna D. 1995. *Participating in Explanatory Dialogues*. Cambridge, Mass.: MIT Press.

Moulin, B., Irandoust, H., Belanger, M. and Desbordes, G. 2002. "Explanation and Argumentation Capabilities". *Artificial Intelligence Review* 17: 169–222.

Nehamas, Alexander 1990. "Eristic, Antilogic, Sophistic, Dialectic: Plato's Demarcation of Philosophy from Sophistry". *History of Philosophy Quarterly* 7: 3–16.

Oddie, Graham 2007. "Truthlikeness". *Stanford Encyclopedia of Philosophy*, available at: http://plato.stanford.edu/entries/truthlikeness/

Parsons, Simon and. Jennings, Nicholas R 'Negotiation through Argumentation: A Preliminary Report', *Proceedings of the Second International Conference on Multi-Agents Systems*, ed. Mario Tokoro, AAAI Press, Menlo Park, California, 1997, 267–274.

Pascal, Blaise 1966. "Reflections on Geometry and the Art of Persuading" (*De L'Esprit Geometrique*, 1659). In Robert W. Gleason (ed.), *The Essential Pascal*, New York: New American Library, 297–327.

Paglieri, Fabio and Castelfranchi, Cristiano 2005. "Arguments as Belief Structures". In David Hitchcock and Daniel Farr (eds.), *The Uses of Argument: Proceedings of a Conference at McMaster University*, Ontario Society for the Study of Argumentation, 356–367.

Peng, Yun and Reggia, James A. 1990. *Abductive Inference Models for Diagnostic Problem-Solving*. New York: Springer-Verlag.

Pereleman, Chaim and Olbrechts-Tyteca, Lucie 1969. *The New Rhetoric: A Treatise on Argumentation*. Notre Dame: University of Notre Dame Press.

Pitt, Jeremy and Mamdani, Abe 1999. "Some Remarks on the Sematics of FIPA's Agent Communication Language". *Autonomous Agents and Multi-Agent Systems* 2: 333–356.

Pitt, Jeremy and Mandami, Abe 2000. "Some Legal Aspects of Inter-Agent Communication". In Frank Dignum and Mark Greaves (eds.), *Issues in Agent Communication*, Berlin: Springer-Verlag, 46–62.

Popper, Karl R. 1963. *Conjectures and Refutations*. London: Routledge.

Prakken, Henry 1995. "From Logic to Dialectics in Legal Argumentation". In *Proceedings of the Fifth International Conference on Artificial Intelligence and Law*, Washington DC: ACM Press, 165–174.

Prakken, Henry 1997. *Logical Tools for Modelling Legal Argument*. Dordrecht: Kluwer.

Prakken, Henry 2001. "Relating Protocols for Dynamic Dispute with Logics for Defeasible Argumentation". *Synthese* 127: 187–219.

Prakken, Henry 2006. "Coherence and Flexibility in Dialogue Games for Argumentation". *Journal of Logic and Computation*. To appear, (current version is technical report UU-CS-2005–021 Institute of Information and Computing Sciences, Utrecht University).

Prakken, Henry 2006. "Formal Systems for Persuasion Dialogue". *The Knowledge Engineering Review* 21: 163–188.

Prakken, Henry 2005. "AI & Law, and Argument Schemes". *Argumentation*,19: 303–320. Available at: http://www.cs.uu.nl/people/henry/

Prakken, Henry and Sartor, Giovanni 1996. "A Dialectical Model of Assessing Conflicting Arguments in legal Reasoning". *Artificial Intelligence and Law* 4: 331–368.

Prakken, Henry and Sartor, Giovanni 1998. "Modelling Reasoning with Precedents in a Formal Dialogue Game". *Artificial Intelligence and Law* 6: 231–287.

Premack, D. and Woodruff, G. 1978. "Does the Chimpanzee Have a Theory of Mind?". *Behavioral and Brain Sciences* 1: 515–526.

Rahwan, Iyad, Moraitis, Pavlos and Reed, Chris 2005. "Preface". In Iyad Rahwan, Pavlos Moraitis and Chris Reed (eds.), *Argumentation in Multi-Agent Systems: First International Workshop, ArgMAS 2004, Revised Selected and Invited Papers*, Berlin: Springer, I-VIII.

Rappaport, Michael B. 1995. "The Ambiguity Rule and Insurance Law". *Georgia Law Review* 30: 171–257.

Reed, Chris 1998. "Dialogue Frames in Agent Communication". In Y. Demazeau (ed.), *Proceedings of the Third International Conference on Multi-Agent Systems*, IEEE Press, 246–253.

Reed, Chris 2006. "Representing Dialogic Argumentation". *Knowledge-Based Systems* 19: 22–31.

Reed, Chris and Rowe, Glenn 2001. "Araucaria: Software for Puzzles in Argument Diagramming and XML". University of Dundee Technical Report.

Reed, Chris and Walton, Douglas 2007. "Argumentation Schemes in Dialogue". *Proceedings of OSSA 07: Ontario Society for the Study of Argumentation*, CD-ROM.

Reiter, Raymond 1987. "Nonmonotonic Reasoning". *Annual Review of Computer Science* 2: 147–186.

Renon, Luis Vega 1998. "Aristotle's *Endoxa* and Plausible Argumentation". *Argumentation* 12: 95–113.

Rescher, Nicholas 1964. *Introduction to Logic*. New York: St. Martin's Press.

Rescher, Nicholas 1976. *Plausible Reasoning*. Assen: Van Gorcum.

Rescher, Nicholas 1977. *Dialectics*. Albany: State University of New York Press.

Rigotti, Eddo 2007. "Relevance of Context Bound *Loci* to Topical Potential in the Argumentation Stage". To appear.

Rigotti, Eddo and Rocci, Andrea 2006. "Towards a Definition of Communication Context". In M. Colombetti (ed.), *Communication sciences as a multidisciplinary enterprise. Special issue of Studies in Communication Sciences*, to appear

Ristau, C. 1991. "Aspects of the Cognitive Ethology of an Injury-Feigning Bird, the Piping Plover". In C. Ristau, (ed.), *Cognitive Ethology: The Minds of Other Animals*, Hillsdale: Erlbaum, 91–126.

Robinson, Richard 1962. *Plato's Earlier Dialectic*, 2nd ed.. Oxford: Clarendon Press, 1962.

Robinson, T. M. 1979. *Contrasting Arguments: An Edition of the Dissoi Logoi*. New York: Arno Press.

Rocci, Andrea 2006. "Pragmatic Inference and Argumentation in Intercultural Communication". *Intercultural Pragmatics* 3–4: 409–442.

Russell, Stuart J. and Norvig, Peter 1995. *Artificial Intelligence: A Modern Approach*. Upper Saddle River, N. J.: Prentice Hall.

Salmon, Wesley 1970. "Statistical Explanation". In R. Colodny (ed.), *The Nature and Function of Scientific Theories*, Pittsburgh: University of Pittsburgh Press.

Salmon, Wesley 1998. "The Importance of Scientific Understanding". In Wesley Salmon (ed.), *Causality and Explanation,* New York: Oxford University Press, 79–91.

Sandholm, Tuomas and Lesser, Victor 1995. "Issues in Automated Negotiation and Electronic Commerce: Extending the Contract Net Framework". In Victor Lesser (ed.), *Proceedings: First International Conference on Multi-Agent Systems*, Menlo Park, AAAI Press, 328–335.

Schank, Roger C. 1986. *Explanation Patterns: Understanding Mechanically and Creatively*, Hillsdale, New Jersey: Erlbaum..

Schank, Roger C. and Abelson, Robert P. 1977. *Scripts, Plans, Goals and Understanding*. Hillsdale, N. J.: Erlbaum.

Schank, Roger C., Kass, Alex and Riesbeck, Christopher K.1994. *Inside Case-Based Explanation.* Hillsdale, New Jersey: Erlbaum.

Schlegloff, Emanuel A. 1988. "Presequences and Indirection". *Journal of Pragmatics* 12: 55–62.

Schiappa, Edward 1995. "Introduction". In Edward Schiappa (ed.), *Warranting Assent*, Albany: State University of New York Press, ix-xxix.

Scott, A. C., Clancey, W. J., Davis, R. and Shortliffe, E. H. 1977. "Explanation Capabilities of Knowledge-Based Production Systems". In B. G. Buchanan and E. H. Shortliffe (eds.), *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*, Reading: Addison-Wesley.

Scriven, Michael 1962. "Explanations, Predictions and Laws". In H. Feigl and G. Maxwell (eds.), *Minnesota Studies in the Philosophy of Science*, vol. 3, Minneapolis: University of Minnesota Press.

Scriven, Michael 1972. 'The Concept of Comprehension: from Semantics to Software', *Language Comprehension and the Acquisition of Knowledge*, ed. John B. Carroll and Roy O. Freedle, Washington DC, W. H. Winston and Sons, 31–39.

Scriven, Michael 1976. *Reasoning*. New York: McGraw-Hill.

Scriven, Michael 2002. "The Limits of Explication". *Argumentation* 16: 47–57.

Searle, John 1969. *Speech Acts*. Cambridge: Cambridge University Press.

Searle, John 2001. *Rationality in Action*. Cambridge: The MIT Press.

Sedley, David 1998. "Dialectical School". In *Routledge Encyclopedia of Philosophy*, Version 1.0, London: Routledge.

Seeskin, Kenneth 1987. *Dialogue and Discovery: A Study in Socratic Method*. Albany: State University of New York Press.

Segerberg, Krister 1984. "Towards an Exact Philosophy of Action". *Topoi* 3: 75–83.

Segerberg, Krister 1985. "Routines". *Synthese* 65: 185–210.

Sextus Empiricus, 1933. *Against the Logicians*, Vol. II of the Loeb Classical Library Works. Cambridge: Harvard University Press.

Shannon, Claude E.and Weaver, Warren 1972. *The Mathematical Theory of Communication*. Urbana: University of Illinois Press.

Sidgwick, Alfred 1893. *The Process of Argument*. London: Adam and Charles Black.

Sidgwick, Alfred 1901. *The Use of Words in Reasoning*. London: Adam and Charles Black.

Sidgwick, Alfred 1910. *The Application of Logic*. London: MacMillan and Co.

Siegel, Harvey and Biro, John 1997. "Epistemic Normativity, Argumentation, and Fallacies". *Argumentation* 11: 277–292.

Sierra, Carles, Jennings, Nicholas R., Noriega, P. and Parsons, Simon 1997. "A Framework for Argumentation-Based Negotiation". In *Proceedings of the Fourth International Workshop on Agent Theories, Architectures and Languages*, Rhode Island, 167–182.

Silverman, Barry G. 1992. *Critiquing Human Error: A Knowledge Based Human-Computer Collaboration Approach*. London: Academic Press.

Singh, Munindar P. 1998. "Agent Communication Languages: Rethinking the Principles". *Computer* 31: 425–445.

Singh, Munindar P. 1999. "A Semantics for Speech Acts". *Annals of Mathematics and Artificial Intelligence* 8: 47–71.

Singh, Munindar P. 2000. "A Social Semantics for Agent Communication Languages". In Frank Dignum and Mark Greaves (eds.), *Issues in Agent Communication*, Berlin: Springer-Verlag, 31–45.

Singh, Push, Lin, Thomas, Mueller, Erik T., Lim, Grace, Perkins, Travell and Wan Li Zhu, 2002. "Open Mind and Common Sense: Knowledge Acquisition from the General Public". In *Proceedings of the First International Conference on Ontologies, Databases, and Applications of Semantics for Large Scale Information Systems*, lecture notes in Computer Science, Heidelberg: Springer-Verlag, 1223–1227.

Smith, Ira A., Cohen, Philip R., Bradshaw, Jeffrey M., Greaves, J. Mark and Holmback, Heather 1998. "Designing Conversation Policies Using Joint Intention Theory". In *Proceedings of the Third International Conference on Multi-Agent Systems (ICMAS-98)*, Paris, France: IEEE Computer Society, 269–276.

Smith, Robin 1993. "Aristotle on the Uses of Dialectic". *Synthese* 96: 335–358.

Snoeck Henkemans, Francisca 1992. *Analyzing Complex Argumentation: The Reconstruction of Multiple and Coordinatively Compound Argumentation in a Critical Discussion*. Amesterdam: SICSAT.

Staniford, G., Bench-Capon, T.J.M. and Dunne, P.E.S. 1993. "Cooperative Dialogues with the Support of Autonomous Agents". In M. Huhns, M. P. Papazogolou, G. Schlageter (eds.), *Proceedings of ICICIS'93 International Conference on Intelligent and Cooperative Information Systems, Rotterdam*. Los Alamitos, California: IEEE Computer Society Press, 144–152.

Stevenson, Charles L. 1944. *Ethics and Language.* New Haven: Yale University Press.

Stump, Eleonore 1989. Dialectic and its Place in the Development of Medieval Logic. Ithaca: Cornell University Press.

Tarski, Alfred 1944. "The Semantic Conception of Truth and the Foundations of Semantics". Philosophy and Phenomenological Research 4, 341–376. Available at http://www.ditext.com/tarski/tarski.html

Tarski, Alfred 1956. "The Concept of Truth in Formalized Languages". In Alfred Tarski, Logic, Semantics, Metamathematics, Oxford: Clarendon Press, 152–278.

Timmerman, David M. 1993. "Ancient Greek Origins of Argumentation Theory: Plato's Transformation of Dialegesthai to Dialectic". *Argumentation and Advocacy* 29: 116–123.

Tindale, Christopher W. 1999. *Acts of Arguing: A Rhetorical Model of Argument*. Albany: State University of New York Press.

van Eemeren, Frans H. and Grootendorst, Rob 1984. *Speech Acts in Communicative Discussions*. Dordrecht: Foris.

van Eemeren, Frans H. and Grootendorst, Rob 1987. "Fallacies in Pragma-Dialectical Perspective". *Argumentation* 1: 283–301.

van Eemeren, Frans H. and Grootendorst, Rob 1992. *Argumentation, Communication and Fallacies*. Hillsdale, N. J.: Erlbaum.

van Eemeren Frans H. and Gootendorst, Rob 1995. "Perelman and the Fallacies". *Philosophy and Rhetoric* 28: 122–133.

van Eemeren, Frans H., Grootendorst, Rob, Snoek Henkemans, Francisca et al., 1996. *Fundamentals of Argumentation Theory*. Mahwah, New Jersey: Erlbaum.

van Eemeren, Frans. H. and Houtlosser, Peter 2000. "Rhetorical Analysis Within a Pragma-Dialectical Framework". *Argumentation* 14: 293–305.

van Eemeren, Frans H. and Houlosser, Peter 2002. "Strategic Maneuvering with the Burden of Proof". In Frans H. van Eemeren et al. (eds.), *Advances in Pragma-Dialectics*, Amsterdam: SicSat, 13–28.

van Fraassen, Bas C. 1993. "The Pragmatics of Explanation". In David-Hillel Ruben (ed.), *Explanation*, Oxford: Oxford University press, 275–309.

van Laar, Jan Albert 2003. *The Dialectic of Ambiguity*. Ph.D. Thesis, University of Groningen.

Verheij, Bart 1996. *Rules, Reasons and Arguments: Formal Studies of Argumentation and Defeat*. Doctoral Dissertation, University of Maastricht.

Verheij, Bart 1999. "Automated Argument Assistance for Lawyers". In *The Seventh International Conference on Artificial Intelligence and Law: Proceedings of the Conference*, New York, N. Y.: ACM, 43–52. Available on http://www.ai.rug.nl/~verheij/.

Verheij, Bart 2000. "Dialectical Argumentation as a Heuristic for Courtroom Decision Making". Web page of Bart Verheij, Department of Metajuridica, Universiteit Maastricht, http://www.ai.rug.nl/~verheij/.

Verheij, Bart 2005. *Virtual Arguments*: *On the Design of Argument Assistants for Lawyers and Other Arguers*. The Hague: Asser Press.

von Wright, Georg Henrik 1971. *Explanation and Understanding*. Ithaca, New York: Cornell University Press.

von Wright, Georg Henrik 1997. "Explanation and Understanding of Actions". In Ghita Holmstrom-Hintikka and Raimo Tuomela (eds.), *Contemporary Action Theory*, Dordrecht: Kluwer, 1–20.

Vreeswijk, Gerard A. W. 1997. "Abstract Argumentation Systems". *Artificial Intelligence* 90: 225–279.

Wagenaar, Willem A., van Koppen, Peter J. and Crombag, Hans F. M. 1993. *Anchored Narratives: The Psychology of Criminal Evidence*. Hertfordshire: Harvester Wheatsheaf.

Walton, Douglas 1989a. *Informal Logic*. New York: Cambridge University Press.

Walton, Douglas 1989b. *Question-Reply Argumentation*. New York: Greenwood Press.

Walton, Douglas 1990. *Practical Reasoning*. Savage, Maryland: Rowman and Littlefield.

Douglas, Walton 1990. "What is Reasoning? What is an Argument?". *Journal of Philosophy* 87: 399–419.

Walton, Douglas 1991. *Begging the Question*. New York: Greenwood Press.

Walton, Douglas 1992. *Slippery Slope Arguments*. Oxford: Clarendon Press, 1992. Reprinted by Vale Press, Newport News, Virginia, 1999.

Walton, Douglas 1995. *A Pragmatic Theory of Fallacy*. Tuscaloosa: University of Alabama Press.

Walton, Douglas 1996. *Argumentation Schemes for Presumptive Reasoning*. Mahwah, New Jersey: Erlbaum.

Walton, Douglas 1996a. *Argument Structure: A Pragmatic Theory.* Toronto: University of Toronto Press.

Walton, Douglas 1996b.*Fallacies Arising from Ambiguity*. Dordrecht: Kluwer.

Walton, Douglas 1997. *Appeal to Expert Opinion*. University Park. Pa.: Penn State Press.

Walton, Douglas 1998. *Ad Hominem Arguments*. Tuscaloosa: University of Alabama Press.

Walton, Douglas 1998a. *The New Dialectic: Conversational Contexts of Argument*. Toronto: University of Toronto Press.

Walton, Douglas 1999. "Profiles of Dialogue for Arguments from Ignorance". *Argumentation* 13: 53–71.

Walton, Douglas 2004. *Abductive Reasoning*. Tuscaloosa: University of Alabama Press, 2004.

Walton, Douglas 2005. *Argumentation Methods for Artificial Intelligence in Law*. Berlin: Springer (Lecture Notes in Artificial Intelligence Series).

Walton, Douglas 2006. *Fundamentals of Critical Argumentation*. Cambridge University Press: New York.

Walton, Douglas 2006a. "How to Make and Defend a Proposal in Deliberation Dialogue". *Artificial Intelligence and Law* 14: 177–239.

Walton, Douglas 2007. "Dialogical Models of Explanation", *Explanation-Aware Computing: Papers from the 2007 AAAI Workshop*, Association for the Advancement of Artificial Intelligence, Technical Report WS-07-06, ed. Thomas Roth-Berghofer, Stefan Schulz and David B. Leake, Menlo Park California, AAAI Press, 1–9.

Walton, Douglas and Krabbe, Erik C. W. 1995. *Commitment in Dialogue: Basic Concepts of Interpersonal Reasoning*. Albany: State University of New York Press.

Wells, Simon and Reed, Chris 2006. "Knowing when to Bargain: the Roles of Negotiation and Persuasion in Dialogue". In *Proceedings of the ECAI Workshop on Computational Models of Natural Argument*, Riva del Garda, Italy: Information Technology Center.

Warnick, Barbara 2000. "Two Systems of Invention: The Topics in the Rhetoric and The New Rhetoric". In Alan G. Gross and Arthur E. Walzer (eds.), *Rereading Aristotle's Rhetoric*, Carbondale: Southern Illinois University Press: 107–129.

Whately, Richard 1863. *Elements of Rhetoric*, 7th ed. London: Parker Son and Bourn.

Wick, M. R. and Thompson, W. B. 1992. "Reconstructive Expert System Explanation". *Artificial Intelligence* 54: 33–70.

Wilensky, Robert 1983. *Planning and Understanding: A Computational Approach to Human Reasoning*. Reading, Mass.: Addison-Wesley.

Williams, M. Lee and Goss, Blaine 1975. "Equivocation: Character Insurance". *Human Communication Research* 1: 265–270.

Woods, John and Walton, Douglas 1978. "Arresting Circles in Formal Dialogues". *Journal of Philosophical Logic* 7: 73–90.

Wooldridge, Michael 2000. *Reasoning about Rational Agents*. Cambridge, Mass.: The MIT Press.

Wooldridge, Michael 2002. *An Introduction to MultiAgent Systems*. Chichester: Wiley.

Wooldridge, Michael and Jennings, Nicholas R. 1995. "Intelligent Agents: Theory and Practice". *The Knowledge Engineering Review* 10: 115–152.

Wooldridge, Michael, McBurney, Peter and Parsons, Simon 2005. "On the Meta-Logic of Arguments". In *Proceedings of the Fourth International Joint Conference on Autonomous Agents and Multi-Agent Systems*, Utrecht.

Zarefsky, David 1990. *Lincoln, Douglas and Slavery: In the Crucible of Public Debate*. Chicago: University of Chicago press.

Zarefsky, David, Miller-Tutzauer, Carol and Tutzauer, Frank E. 1984. "Reagan's Safety Net for the Truly Needy: The Rhetorical Uses of Definition". *Central States Speech Journal,* 35: 113–119

Index

In the series *Controversies* the following titles have been published thus far or are scheduled for publication: